

## POLICY

### European research

The European Research Council (ERC), a granting agency for basic research, should remain part of the European Commission, according to a report published on 12 July by a commission task force. But the council's governance should change to allow the ERC to side-step burdensome commission bureaucracy, such as demands for detailed time sheets and auditing. The task force recommends that the ERC appoints a near-full-time president, based in Brussels, and makes its governing body, the Scientific Council, more powerful. The ERC launched in 2007 and has so far distributed more than €3 billion (US\$4.2 billion) to top researchers around Europe.

### Efficient ships

The International Maritime Organization (IMO) has approved internationally binding energy-efficiency standards for ships — the world's first such regulations for any industry. Adopted during a meeting at IMO headquarters in London on 15–19 July, the regulations require a 10% increase in energy efficiency for ships of more than 400 tonnes built between 2015 and 2019, rising to 30% for ships built after 2024. Environmental groups expressed cautious optimism, but criticized loopholes that could allow developing countries to delay implementation by several years.

### Fisheries reform

The European Commission on 13 July published plans to overhaul Europe's fishing industry. The commission wants to cut catch quotas to

sustainable levels by 2015. It also hopes to ban the practice of throwing away the unwanted by-catch. Critics worry that European member states will not be required to heed scientists' advice on catch quotas. See [go.nature.com/plwhgk](http://go.nature.com/plwhgk) for more.

### Carbon capture

A keenly watched US project to capture and bury carbon dioxide has been put on hold. American Electric Power of Columbus, Ohio, had been granted US\$334 million from the Department of Energy to help capture carbon dioxide from its Mountaineer coal-fired power plant in New Haven, West Virginia, and to sequester around 1.5 million tonnes of

gas underground each year. On 14 July the company suspended the project, citing the weak economy and the uncertain status of US climate policy.

### Nuclear-free Japan

Japan's prime minister Naoto Kan has urged his country to cut its nuclear dependency in the wake of the Fukushima disaster. In a televised news conference on 13 July, he said Japan's aim should be to manage without nuclear power — but didn't say when.

### US nuclear safety

A task force appointed by the US Nuclear Regulatory Commission has suggested a range of improvements to safety requirements for the

in Arlington, Virginia, released the image on 14 July. The group announced that the species had been rediscovered in Malaysia late last year by a team led by Indraneil Das from Malaysia Sarawak University (UNIMAS).



## Technicolour toad caught on camera

The striking Bornean rainbow toad (*Ansonia latidisca*; pictured) has been seen for the first time since the 1920s — and this time, it has been caught on camera. The non-profit group Conservation International, headquartered

country's nuclear reactors — targeting everything from power outages, flooding and earthquakes to issues of emergency preparedness and the cooling of spent-fuel storage pools. The commission must now decide how, whether and when to implement the recommendations, which were issued on 13 July. See [go.nature.com/15x8r2](http://go.nature.com/15x8r2) for more.

### Tuition hikes

Universities in California have raised tuition fees — again — in response to more cuts in state funds. On 12 July, the 23-campus California State University system increased fees by 12% for this autumn, on top of an existing 10% rise that had been approved last

NASA/JPL-CALTECH/UCLA/MPS/DLR/IDA  
November. Two days later, the 10-campus University of California agreed a 9.6% increase, adding to an 8% hike that it had already approved, also last November. The university systems are each facing \$650-million cuts in state funds this year (see *Nature* 475, 10; 2011).

## Whaling failure

This year's meeting of the International Whaling Commission (IWC) ended acrimoniously after attendees walked out of the conference in St Helier, Jersey, on its final day, 14 July. The stumbling block was a years-old proposal to create a sanctuary for whales in the south Atlantic. Japan and other pro-whaling nations said that they didn't want to vote on the issue, because "reverting to voting could be harmful to the constructive dialogue and atmosphere in the Commission", the IWC reported. The countries' delegates then left, and the vote was shelved until next year's meeting.

## RESEARCH

## Asteroid tour

NASA's Dawn spacecraft has entered orbit around Vesta, which sits in the main asteroid belt between Mars and Jupiter (pictured). At 500 kilometres in width, it is the second-largest



asteroid in the Solar System, with a geological structure apparently like that of a miniature Earth: a metallic core, and rocky mantle and crust (see *Nature* 475, 147–148; 2011). Dawn, which launched in September 2007, will spend a year touring Vesta before firing up its ion engines to reach the largest asteroid, Ceres, in 2015.

## Stem-cell trials

The world's second and third clinical trials of therapies using human embryonic stem cells have begun. Advanced Cell Technology, based in Santa Monica, California, said last week that the first subjects in each of its trials were treated on 12 July. One trial is for a rare cause of blindness — Stargardt disease, which usually strikes in childhood. The other is for a more common cause — age-related macular degeneration. Both are safety studies. The first stem-cell trial, a safety study for a

spinal-cord injury therapy by Geron of Menlo Park, California, started in October 2010. See [go.nature.com/yfjtwk](http://go.nature.com/yfjtwk) for more.

## Radio record

A Russian radio telescope that launched from Baikonur, Kazakhstan, on 18 July, is orbiting Earth at about 90% of the distance to the Moon, reports Roscosmos, the Russian space agency. That distance allows the Spectrum-R instrument, also known as RadioAstron, to create extremely high-resolution images by combining its own observations with those from telescopes based on Earth. The effective size of the 'dish' made by the combination will be more than ten times that created in the same way by the Japanese radio telescope HALCA, which orbited Earth from 1997 to 2005.

## HIV prevention

Giving antiretroviral drugs to HIV-negative people is an effective way of preventing HIV infection, two teams behind clinical trials in Africa reported at the meeting of the International AIDS society in Rome on 17–20 July. The TDF2 trial on more than 1,200 uninfected heterosexuals in Botswana was found to cut the risk of HIV infection by up to 78%. And the Partners

## COMING UP

### 26 JULY

The US National Institutes of Health discovers how much Republicans in the US House of Representatives want to lop from its 2012 budget. In recent weeks, heavy cuts have been proposed for a number of US science agencies.

PrEP trial, on around 4,750 heterosexuals in Kenya and Uganda, reduced the chance of infection by up to 73%. See [go.nature.com/bdfsci](http://go.nature.com/bdfsci) for more.

## GM crop vandals

Three Greenpeace activists broke into a scientific farm near Canberra overnight on 14 July and destroyed a crop of genetically modified wheat. The crop — engineered to have a lower glycaemic index, so releasing glucose more gradually, and a higher fibre content — was a field trial by Australia's national science agency, the CSIRO. See [go.nature.com/ie5wv4](http://go.nature.com/ie5wv4) for more.

## PEOPLE

## Hawass fired

Zahi Hawass, Egypt's antiquities minister, has lost his post. The move is part of a cabinet reshuffle to appease critics of the current government, which has retained many officials who were part of the regime of former president Hosni Mubarak. Hawass, a prominent archaeologist, was appointed minister of state for antiquities in the dying days of Mubarak's tenure. He quit in March — but was reappointed within a month. A permanent replacement had not been confirmed as *Nature* went to press.

➔ [NATURE.COM](http://NATURE.COM)

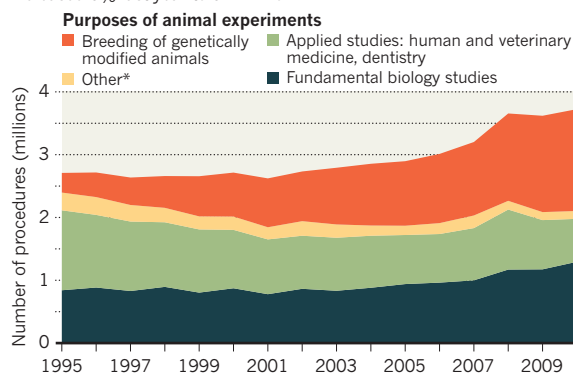
For daily news updates see:  
[www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

The number of genetically modified animals — particularly mice — being bred for research in the United Kingdom rose by 6% from 2009 to 2010, continuing on from last year's rise of 10%. The increase was largely responsible for the greater number of procedures using animals in UK research reported by the Home Office on 13 July (see chart). Experiments on zebrafish also rose, although fewer cats, dogs and horses were used than last year. See [go.nature.com/cegyina](http://go.nature.com/cegyina) for more.

## ANIMAL USE RISES IN UK RESEARCH

The number of procedures performed on animals in UK research increased 3% last year to 3.7 million.



# THIS WEEK

## EDITORIALS

**E-SPECIES** Zoologists should follow botanists into the digital age **p.424**

**WORLD VIEW** Meet the world's only blogging stem-cell scientist **p.425**

**MAN BITES DOG** Ancient canines more than best friends **p.427**

## The legacy of Doctor Moreau

*Regulators must look past visceral disgust about human–animal hybrids. Strict but sensible rules are needed for research on hybrid embryos and chimaeric animals that could produce therapies.*

The science-fiction author H. G. Wells coined the term humanized animals in his 1896 novel *The Island of Doctor Moreau*. The book invited readers to consider the ethical limits of curiosity-driven research and to ponder the moral value of the distinction between humans and animals. The book's evil protagonist creates, through a vaguely defined process of 'vivisection', a colony of half-human 'beast folk', unhappy in themselves and frightening to others.

Dr Moreau's humanized animals evoke visceral disgust. Thankfully, more than a century later, they remain science fiction. However, the ethical dilemmas presented by Wells do not.

Innumerable mice and other animals have been engineered in past decades to express a human gene and model specific aspects of human disease. They rarely inspire disgust, because they still resemble their own species. But further advances in genetic and stem-cell technologies mean that researchers could, in theory, create animals with quintessentially human characteristics or behaviours. The sight of an animal with shiny, furless 'human' skin, for example — exceptionally useful for research into skin disorders — could evoke disgust similar to that created by Moreau's beast folk, even though the animal itself might be perfectly comfortable. One of the biggest horrors — although technically unlikely — could be a self-aware monkey, a creature with human thought trapped in the body of an animal, unable to express itself.

Prompted by the possibilities, scientists around the world have begun to discuss the ethical consequences of taking to extremes the frontier technologies that allow mixing of species. These include the introduction of human stem cells into animals, where they could integrate into the animal's body; or the formation of hybrid or chimaeric embryos that mix the DNA of humans and animals.

The UK Academy of Medical Sciences in London has now produced a comprehensive report on the subject (see page 438 and page 448). The document is likely to lead to pioneering legislation specifically geared towards regulating research on animals containing human material. This is a timely and important step: timely because little truly controversial research in this area has yet been done, so both public debate and scientific research can take place in a peaceful environment; important because instinctive revulsion should not automatically block future research that will undoubtedly pave the way for therapies for currently incurable diseases.

The report clearly identifies techniques that cannot yet be used ethically, including extensive humanization of the monkey brain or the development of embryos that mix DNA from humans and non-human primates. It distinguishes them from procedures such as the creation of transgenic mice that bear human genes, which the academy says require no oversight beyond the strict controls that already apply.

The academy says that further oversight of experiments that occupy the middle ground — perhaps the human-skinned animal — should assess, on a case-by-case basis, whether the benefits to understanding or to medicine from a particular research project outweigh the potential

suffering of the animal involved, its cage mates or its carers.

The UK government commissioned the report and is likely to adopt its conclusions. In doing so, it will reinforce Britain's reputation as an attractive research environment, strictly controlled but without unwarranted hindrances. The country has some of the world's most stringent laws on the welfare of research animals, but also some of the most rational regulations for research using human embryonic stem cells. It

**"Advances in genetic and stem-cell technologies could, in theory, create animals with quintessentially human characteristics."**

allows the creation of hybrid embryos that are predominantly human — forbidden in many countries — as long as they are destroyed before they develop beyond the two-cell stage. Now the country seems ready to regulate hybrid embryos that are mainly animal, as well as chimaeric animals.

The United Kingdom knows that this will give it an advantage in reaping medical benefits — and that proactive legislation offers protection against future calls for outright bans, should public anxiety grow. Potential therapies using human stem cells to replace damaged organs or tissue must first be tested in animals. Chimaeric animals with human brain material might be useful. For example, they could help to establish how the normal human brain develops and functions, and what goes wrong in neurodevelopmental disorders such as schizophrenia.

The ethical questions raised by H. G. Wells are as valid today as they ever were. But as facts and fiction converge, the answers have become more complex. ■

## Heart of the matter

*The Heartland Institute's climate conference reveals the motives of global-warming sceptics.*

It would be easy for scientists to ignore the Heartland Institute's climate conferences. They are curious affairs designed to gather and share contrarian views, in which science is secondary to wild accusations and political propaganda. They are easy to lampoon — delegates at the latest meeting of the Chicago-based institute in Washington DC earlier this month could pick up primers on the libertarian writings of Russian–American novelist Ayn Rand, who developed the philosophical theory of objectivism, and postcards depicting former US vice-president Al Gore as a fire-breathing demon. And they are predictable, with environmentalists often portrayed as the latest incarnation of a persistent communist plot. "Green on the outside, red on the inside,"



said one display. “Smash the watermelons!”

So why does *Nature* this week devote two pages to such absurdities? We now have more than two decades of evidence that closing our eyes will not make the climate sceptics go away. Instead, in the United States at least, they have cemented their propaganda into a broader agenda that pits conservatives of various stripes against almost any form of government regulation. The sceptics like to present the battlefield as science, but, as the News Feature on page 440 makes clear, the fight is, in fact, a violent collision of world views.

Does the following sound familiar? “They distort science, ignore reality and will not tolerate opinions or facts that conflict with their beliefs.” “Cynical manipulators or simple pawns, their purpose is only to keep funds flowing to a corrupt few who profit from the status quo.” Those are the kinds of words scientists use, often correctly, to describe the sceptics, many of whom would have the financial interests of today continue their dominance tomorrow. Yet this is also how sceptics characterize climate scientists, whose careers and reputations they claim are intertwined with protecting the science of anthropogenic global warming.

To address this conflict might be seen as lending respectability to the spurious claims made by sceptics against respected scientists and robust science. So, let’s be clear: *Nature* is not endorsing the Heartland Institute as a serious voice on climate science. Instead, the News Feature is intended to offer researchers outside climate science a window into the motives and tactics of those who have set themselves up as such a voice. (Those inside climate science, of course, are all too aware of these already.)

Despite criticizing climate scientists for being overconfident about their data, models and theories, the Heartland Institute proclaims a

conspicuous confidence in single studies and grand interpretations. A 2009 report by the Nongovernmental International Panel on Climate Change, which the institute supports, is well sourced and based on scientific papers. Yet it makes many bold assertions that are often

**“Nature is not endorsing the Heartland Institute as a serious voice on climate science.”**

questionable or misleading, and do not highlight the uncertainties. Many climate sceptics seem to review scientific data and studies not as scientists but as attorneys, magnifying doubts and treating incomplete explanations as falsehoods rather than signs of progress towards the truth. As the News Feature points out, although the sceptics feel

that they have already won the political battle in the United States, their attacks on science will continue.

Scientists can only carry on with their work, addressing legitimate questions as they arise and challenging misinformation. Many climate scientists have already tried to engage with their critics, as they did at the Heartland event. The difference, of course, is motive. Scientists work to fill the gaps in human knowledge and to build a theory that can explain observations of the world. Climate sceptics revel in such gaps, sometimes long after they have been filled.

It is scientists, not sceptics, who are most willing to consider explanations that conflict with their own. And far from quashing dissent, it is the scientists, not the sceptics, who do most to acknowledge gaps in their studies and point out the limitations of their data — which is where sceptics get much of the mud they fling at the scientists. By contrast, the Heartland Institute and its ilk are not trying to build a theory of anything. They have set the bar much lower, and are happy muddying the waters. ■

# Origin of species

*Zoologists should follow botanists in allowing online-only announcements of new species.*

**I**da had a difficult birth. This remarkably well preserved primate fossil was introduced to the world in May 2009, in a description in the online journal *PLoS ONE*. Arguments raged over the evolutionary importance of the unquestionably photogenic new species *Darwinius masillae* and the role of a television company in its unveiling; a more technical criticism came from taxonomy specialists. Under the International Code of Zoological Nomenclature, publication of species names must be in a ‘durable medium’ — that is, on paper or CD-ROM, which must be made available at specific libraries. The online-only naming of Ida as a new species was therefore illegitimate. The journal rushed out a correction stating that a “separate print-only edition is available” to fulfil the requirements and ensure that Ida really was *D. masillae* (J. L. Franzen *et al.* *PLoS ONE* 5, e5723; 2009). But the underlying problem remained: if zoologists and botanists wished to publish their findings in online journals, they would still have to physically distribute print copies to suitable repositories.

Last year, Sandra Knapp, a botanist at the Natural History Museum in London, got around the plant-science version of the same rules, the International Code of Botanical Nomenclature. She described four new plants, again in *PLoS ONE* (S. Knapp *PLoS ONE* 5, e10502; 2010), printed out hard copies of the online paper, and posted them to libraries.

At a meeting in Melbourne, Australia, last week, the International Botanical Congress took the first steps to end this situation and bring botany into the electronic age (see *Nature* <http://dx.doi.org/news.2011.428>; 2011). The congress’s nomenclature section, chaired by Knapp, voted to amend the code to allow purely

electronic publication. They also agreed to abandon the need for a Latin description, although the names themselves must still be Latinized. The wider congress must endorse the changes before they can take effect, but it is expected to do so in time for the new rules to apply from January next year.

Now, zoologists should follow suit. More than ten years have passed since the first creatures — fossilized microorganisms called protists — were described in an electronic journal (D. B. Scott *et al.* *Palaeontol. Electron.* 3; 2000; see [go.nature.com/pt1cmz](http://go.nature.com/pt1cmz)), with physical copies sent to libraries to meet the criteria. *Nature* greeted the news at the time with a headline that now seems premature: ‘Online naming of species opens digital age for taxonomy’ (see *Nature* 408, 278; 2000).

Yet, despite much discussion in the scientific community and numerous articles published on the topic, both in print and online, the rules governing zoologists remain as strict as ever.

Researchers pushing for online publications to be given equal status insist that electronic copies can now be considered a permanent record. They say that widening the number of journals that can publish discoveries will benefit the field, and that online-only journals often publish faster than traditional print publications.

At this point, it seems that there is little reason to continue to demand paper on a shelf to make a species name official.

Sometimes, taxonomists — and scientific publishers — resist change and the adoption of new technology. There are often good reasons for this. Proper rules are important and taxonomic anarchy would ruin science. Slavishly embracing every new technology as it appeared would be a disaster, and a demonstrably robust archive is invaluable — just ask a historian.

But electronic publication has already altered the face of publishing, and it will continue to do so. Taxonomy and publishing are likely

**➔ NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xbhunq](http://go.nature.com/xbhunq)

always to lag behind technological progress; in fact, both already have to run faster and faster just to keep up. Still, botanists are about to narrow the gap slightly, and zoologists should pick up the pace. ■





## My year as a stem-cell blogger

Paul Knoepfler explains why he joined the ranks of the blogosphere, and why you should too.

The final week of October 2009 started badly for me, when *Nature* closed its stem-cell blog The Niche — one of my favourites as it covered my field. A few days later, events took a more serious turn for the worse when — out of the blue — I was diagnosed with aggressive prostate cancer. I was 42. Blogs and blogging were suddenly the last thing on my mind.

More than a year and a half later, I am in remission and doing great. The battle against the cancer was the most difficult of my life and I still worry that it may come back. But the experience also had positive effects. For one, I still missed The Niche, and assumed that others did too. Once I recovered, I found the courage to start a replacement. After all, how hard can blogging be when compared with facing cancer? A year on, it has been a remarkable experience.

Compared with just two years ago, blogging is much more accepted as a means of communication by scientists, many of whom read blogs even if they do not write them. Yet, in my experience, it remains relatively rare for faculty-level academic biological researchers to blog regularly, particularly in controversial fields such as stem-cell research. As far as I know, I am the only stem-cell faculty scientist in the world who runs a regularly updated blog. (If I am wrong then please get in touch to say hello!)

Why is this? Other scientists in academia tell me they worry that blogging would damage their careers. Specifically, they fear that colleagues would view them as amateurs, 'wasting time' on blogging, which could reduce their chances of achieving tenure. They fear the wrath of others in the field should they post the 'wrong' thing on their blog, and they worry about payback in negative grant and paper reviews. Some are concerned about attracting unruly and insulting readers' comments.

Are they right to be fearful? In an entire year of blogging I have had to censor just six inflammatory or defamatory comments. Despite my blog taking on the anti-stem-cell community in the United States and the misinformation its members peddle, such as the meme that adult stem cells are a panacea that make embryonic stem cells redundant, I have received remarkably few personal attacks from them. I am grateful for that, if puzzled.

This is certainly not because my blog goes unnoticed. True, I started with just five readers a day, but one year later, traffic has increased more than 30-fold and continues to rise. The blog averages 150 visitors a day and sometimes up to 500 a day, made up of a veritable *Who's Who* in stem-cell science, and beyond. How do I know? Senior figures in the field tell me in confidence that they read and enjoy the blog, although none has publicly

contributed on it — perhaps a sign that there is still a way to go before scientists stop being nervous about blogs.

My audience extends beyond academics, and discussion on the blog has catalysed links between academic scientists, biotechnology companies, big pharma, government officials, funding agencies, investors, teachers, reporters, students and patient advocates.

The blog has also helped to coordinate political efforts to advance stem-cell research through campaigns that put pressure on US and Californian officials to vote pro-science and pro-stem cells. I find self-publishing on my blog a liberating break from the tedious and frustrating grind of peer review, and it has encouraged me to write more for general audiences.

There has been some negative feedback, usually expressed privately rather than on the blog itself. Some critics cautioned that I might anger 'the wrong people' in academia or at funding agencies. Others were more direct with their disapproval of some of my most popular posts — usually those that mentioned specific funding agencies or companies by name — with the implied threat that I would see papers or grant applications rejected. Some who disliked my outspokenness insisted I shut down the blog. But don't let this put you off — the threats came from a very small number of people and have (so far) been toothless. My lab is fine in terms of funding and publishing, and I recently got tenure.

Here are some tips for beginners. Start slowly; wait a day after writing and reread your draft before posting. Try to avoid discussing your own institution, and critique papers or theories in the field in a constructive manner. It is important

that you include your own opinions, but do not use your blog to broadcast your opinions about issues that are unrelated to science.

Update your blog regularly, because readers will not visit blogs that they perceive as boring or 'old news'. Read and comment on other blogs, which will lead people to yours. Get a Twitter account to promote it and dabble with search-engine optimization. And do tell your colleagues about your blog.

Savvy scientists must increasingly engage with blogs and social media. A new generation of young researchers has grown up with an ever-present Internet. Publishers have been quicker than academics to react to this new world, but scientists must catch up. Even if you choose not to blog, you can certainly expect that your papers and ideas will increasingly be blogged about. So there it is — blog or be blogged. ■

Paul Knoepfler is an associate professor at the University of California, Davis, School of Medicine and blogs at [www.ipsell.com](http://www.ipsell.com). e-mail: [knoepfler@ucdavis.edu](mailto:knoepfler@ucdavis.edu)

WE STILL HAVE  
**SOME WAY**  
TO GO BEFORE  
**SCIENTISTS**  
STOP BEING  
**NERVOUS**  
ABOUT BLOGS.

**NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/cmvc9v](http://go.nature.com/cmvc9v)

# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## STEM CELLS

### Edited stem cells yield healthy mice

By reprogramming mouse cells into stem cells and purging them of a disease mutation, researchers have spawned disease-free animals.

Tobias Cantz at Hanover Medical School in Germany and his colleagues began by making induced pluripotent stem cells (iPSCs) from the cells of mice lacking the enzyme FAH, which are used as a model of some hereditary liver diseases. The authors then delivered a working version of the *Fah* gene to the iPSCs with a retrovirus. Finally, they used these cells to create embryos and generate mice with working copies of *Fah* and no signs of liver disease.

Human iPSCs might be similarly manipulated to rid them of mutations and then converted into working tissues for transplantation.

*PLoS Biol.* 9, e1001099 (2011)

## MATERIALS

### Shape-shifter forms box in water

Just add warm water, and a specially designed polymer composite can fold into a box or even a flower — on its own.

The material, a hydrogel embedded with carbon nanotubes, was conceived by Ali Javey at the University of California, Berkeley, and his collaborators. At room temperature, the composite is hydrophilic and water is incorporated in its structure. But above a threshold

temperature of around 33°C, it becomes hydrophobic and extrudes water. This reversible transition causes a mechanical strain that can be exploited to create shape-changing actuators. The carbon nanotubes ensure a fast response time.

In 48°C water, the material took 35 seconds to fold into a box (**pictured**). Such fast self-folding materials could be useful for applications

including smart solar devices and tissue engineering.

*Nano. Lett.* <http://dx.doi.org/10.1021/nl201503e> (2011)

## NEUROBIOLOGY

### How blasts hurt the brain

Brain injuries caused by explosions result in the retraction of neural connections and blood-vessel

constriction, reducing oxygen flow to the brain. Researchers have identified proteins and a signalling pathway that link mechanical forces to these harmful changes.

Kit Parker at Harvard University in Cambridge, Massachusetts, and his colleagues simulated blast forces on rat neurons using devices that abruptly stretched the cells. The authors showed that such forces stimulate transmembrane proteins called integrins, which are connected to the cell's internal skeleton. This stimulation overactivates the Rho-ROCK signalling pathway that normally controls cell contraction. Blocking this



M. PAWLITZKI/PHOTOLIBRARY

## GEOCHEMISTRY

### Air pollutants make a comeback

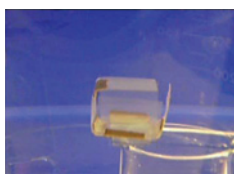
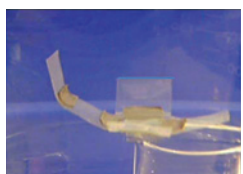
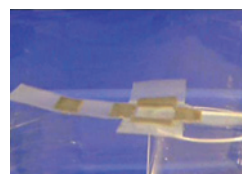
Climate warming is remobilizing toxic pollutants deposited in Arctic ice and sea water.

Long-lived or 'persistent' organic pollutants (POPs), which include chlorinated pesticides and industrial chemicals, travel in the atmosphere to the high north, where they accumulate. They have been regulated for several decades, but sea-ice retreat and rising temperatures seem to be returning some of the more volatile compounds stored in Arctic reservoirs to the atmosphere.

Jianmin Ma and Hayley Hung of Environment

Canada in Toronto and their team analysed the concentrations of POPs measured since 1993 at two Arctic stations. When the effect of regulation was removed, several of the compounds showed increasing atmospheric levels, corresponding with rising Arctic temperatures and decreasing sea-ice cover. A simulation of the impact of climate change on the chemicals' atmospheric abundance confirmed this finding.

*Nature Clim. Change* <http://dx.doi.org/10.1038/nclimate1167> (2011)



X. ZHANG

pathway decreases neuronal injury.

In a separate study, the same group found that similar mechanisms explain why engineered human vascular tissue constricts when subjected to force. The pathway could be a target for preventive drugs, the authors say.

*PLoS ONE* 6, e22899 (2011); *Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1105860108> (2011)

## MATERIALS

## Soft optics from organic gels

Optical devices such as lenses and prisms can be made from organic gels developed by Kana Sureshan and his colleagues at the Indian Institute of Science Education and Research in Thiruvananthapuram.

They used very low concentrations of two different molecules based on the six-carbon carbohydrate mannitol to make a series of gels from hydrocarbon-based solvents such as paraffin oil and pump oil. Shapes formed from these gels were transparent and remained stable for months. The gels could filter ultraviolet light and bend visible light much like glass does. Moreover, optical devices made from the gels were shatter- and scratch-proof, thanks to the gels' ability to fuse together after being cut into pieces.

*Angew. Chem. Int. Edn* <http://dx.doi.org/10.1002/anie.201103584> (2011)

## ADDICTION

## Target for blocking cocaine

The effects of cocaine can be counteracted in mice by stimulating an enigmatic brain receptor, suggesting a possible target for addiction treatment.

The CB<sub>1</sub> receptor in the brain has been implicated in drug addiction, but the role of the CB<sub>2</sub> receptor has not been clear. So Zheng-Xiong Xi and his team at the National Institute on Drug Abuse in Baltimore,

Maryland, gave cocaine-addicted mice a compound that stimulates the CB<sub>2</sub> receptor. The animals self-administered cocaine less frequently, whereas no effect was seen in mice lacking the CB<sub>2</sub> receptor.

The effect occurred after intranasal but not intravenous injection of the compound, suggesting that it involves CB<sub>2</sub> receptors in the brain but not in other parts of the body.

*Nature Neurosci.* <http://dx.doi.org/10.1038/nn.2874> (2011)

## ARCHAEOLOGY

## Another origin of domestic dogs

The discovery of ancient dog bones in France point to a European origin for man's best friend.

Humans domesticated dogs from wild wolves, and genetic and archaeological studies have indicated that domestication could have occurred in East Asia and the Middle East, as well as in Europe. Maud Pionnier-Capitan at the National Museum of Natural History in Paris and her team uncovered dozens of remains of small dogs (**pictured**) at two sites in southern France and another near Paris. The fossils are 11,500–15,500 years old, making them contemporaries of much larger ancient dogs found in Russia. This indicates that dog domestication may have occurred independently in various parts of Europe and Asia.

Furthermore, some of the bones bore cut marks, suggesting that ancient humans ate some of the canines.

*J. Archaeol. Sci.* 38, 2123–2140 (2011)

## CELL BIOLOGY

## Microtubules beat in sync

Certain cells bear tiny hair-like structures called cilia, which beat in wave-like patterns. The

## COMMUNITY CHOICE

The most viewed papers in science

## GENOMICS

## Customer data aid Parkinson's study

**HIGHLY READ**  
on [www.plosgenetics.org](http://www.plosgenetics.org)  
17 June–17 July

Data collected by a personal genetic testing company in California reveal new associations between two genomic areas and a risk of Parkinson's disease.

Chuong Do and Nicholas Eriksson at 23andMe in Mountain View and their colleagues compared the genomic sequences of more than 3,400 people with Parkinson's disease with those from 29,000 23andMe customers who reported not having the disease in an e-mail questionnaire. The authors confirmed 20 previous genetic associations with the disease in addition to revealing the two new links. However, they suggest that many more genetic associations with the disease remain to be found.

*PLoS Genet.* 7, e1002141 (2011)

mechanism underlying this behaviour is unknown, but researchers have succeeded in mimicking it with bundles of protein filaments.

Zvonimir Dogic and his colleagues at Brandeis University in Waltham, Massachusetts, built the *in vitro* system from just three components. They used microtubules (MTs), the filaments that form the backbone of cilia; kinesin motors, the protein complexes that bind to and slide along MTs; and polyethylene glycol, which bundles the MTs together. The authors found that their bundles, when

anchored at one end, beat much less frequently than cilia — which have more than 650 proteins — but synchronize their beating at high density.

The authors suggest that the sliding kinesin motors, which were engineered to bind to two adjacent MTs, move these MTs in opposite directions, resulting in the beating motion.

*Science* 333, 456–459 (2011)

## MICROBIOLOGY

## Bacteria whip around corners

By using their hair-like projections as a slingshot, bacteria can quickly swerve.

*Pseudomonas aeruginosa* bacteria clump together to form biofilms during pneumonia infection and rely on small appendages called pili to attach to surfaces and pull themselves forwards. Gerard Wong at the University of California, Los Angeles, and his group used video microscopy to track the trajectories of the leading and trailing poles of *P. aeruginosa* cells. They propose that the bacteria can also use the pili to change direction. The rapid release of a single tethered pilus provides enough force to spin a bacterium around.

The authors think that this fast slingshot action helps the microbes to move through the viscous maze of polysaccharides they secrete when forming biofilms.

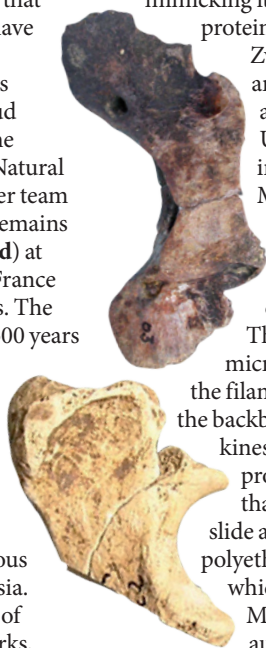
*Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1105073108> (2011)

## NATURE.COM

For the latest research published by Nature visit:

[www.nature.com/latestresearch](http://www.nature.com/latestresearch)

M. PIONNIER-CAPITAN





# NEWS IN FOCUS

**SPACE** NASA's Curiosity rover gets a scenic destination on Mars **p.433**

**PHYSICS** LHC's quest for new particles draws a blank **p.434**

**GENOMICS** Sequencing powerhouse sees a future in computing **p.435**

**US POLITICS** Joe Bast leads the charge against climate science **p.440**



D. DILKOFF/AFP/GETTY IMAGES



Scientific groups have long campaigned on behalf of oppressed researchers and physicians, bolstering international protests such as this candle-light vigil in 2004 for medical workers imprisoned in Libya.

## SCIENTIFIC SOCIETIES

# Human-rights shift under fire

*Critics concerned by scientific bodies' move away from traditional campaigning on behalf of oppressed researchers.*

BY DECLAN BUTLER

Leading scientific societies in the United States are being challenged by critics who say the societies are curtailing their defence of individual cases of persecuted academics.

The critics welcome the fact that the societies, which include the New York Academy of Sciences (NYAS), the American Association

for the Advancement of Science (AAAS) and the American Chemical Society (ACS), are broadening their activities to include scientific diplomacy, and trying to improve access to basic human rights such as clean water, food, health care and education. But they fear that this shift is reducing the emphasis on direct campaigning on behalf of individuals whose human rights have been violated.

Some have also raised concerns about

what they say is a growing reluctance among scientists in general about speaking out on human-rights abuses, in order to avoid interfering with collaborations with scientific powers that have a poor human-rights record, such as China. "I think they should be called out on that," says Robert Quinn, executive director of the Scholars at Risk Network based in New York City, which campaigns on behalf of persecuted academics.

But the societies insist that, far from any retreat from the defence of human rights, their broader and more collaborative approach may help to persuade governments to improve their human-rights records across the board. In addition, the ACS denies that it is reducing its case work on human rights.

*Nature* has learned that in May, the NYAS gave its Committee on Human Rights of Scientists (CHRS) the draft of a new mandate that would curtail its scope to defend individual cases of human-rights abuse, largely restricting it to lending support to cases taken up by other scientific organizations. The NYAS itself would speak out only on exceptional cases.

Ellis Rubinstein, president and chief executive of the NYAS, emphasizes that the mandate has not been finalized. But he argues that the world has changed since the cold war, when most academic human-rights committees were set up, and when "publicly embarrassing" a country was one of the few ways to defend persecuted academics. Today, international science, technology and education partnerships offer opportunities to spur wholesale change in countries, he says, especially for an organization with limited resources, such as the NYAS. The challenge is to find where the NYAS can be most effective, he adds: "You might not be able to help one person, in exchange for trying to help thousands by getting a country to open up."

The CHRS was created in 1978, and made a name for itself when Andrei Sakharov, a Soviet dissident, nuclear-weapons scientist and recipient of the 1975 Nobel Peace Prize, credited it with helping to win his release from internal exile in 1986. It has since engaged in letter writing and other campaigning on behalf of many individual academics, in part to remind oppressive regimes that their handling of such cases is under international scrutiny. Scientists who have suffered abuses, including ►

imprisonment, say that such support is crucial.

Tension has been rising between the CHRS and the NYAS administration since 2007, when the CHRS awarded its annual human-rights prize to two Chinese physicians. China pressured at least one of the recipients, Jiang Yanyong — who was involved in exposing China's cover-up of the SARS epidemic in 2003 — into not travelling to New York to accept the prize. The CHRS wanted to protest against this, but the NYAS administration refused, arguing that such action could damage its relations with China, according to minutes of internal meetings seen by *Nature*. The controversy prompted Svetlana Stone Wachtell, director of the human-rights programme at the NYAS, to resign, and Eugene Chudnovsky, a physicist at the City University of New York, to step down as chair of the CHRS.

Rubinstein argues that the case for protesting to China was less convincing than if the two doctors had been imprisoned. "These people were not being persecuted any more than we sometimes in our own country don't give people visas for a variety of different reasons," he asserts.

## LACKING BACKBONE

The NYAS board has subsequently vetoed all of the CHRS's letters of support for persecuted individuals, "except for letters to Cuba, Iran and Chad", leaving the committee moribund, says Sophie Cook, a member of the CHRS and executive director of the Committee of Concerned Scientists, a human-rights body based in New York City. Rubinstein insists that the academy will not hesitate to speak out on "egregious" individual abuses of human rights, and says that the vetoed letters addressed cases that were not compelling or involved people other than scientists and engineers. Chudnovsky disputes this, arguing that they involved clear cases of human-rights concerns.

The NYAS administration "doesn't want to admit to itself that it is just being cowardly", says Joel Lebowitz, director of the Center for Mathematical Sciences Research at Rutgers University in Piscataway, New Jersey, and a member of the CHRS. Scientific societies "should have a bit more backbone" when it comes to speaking out on human-rights abuses involving countries with which they collaborate.

"That's a reasonable criticism," says Torsten Wiesel, emeritus chairman of the NYAS board of governors and winner of the 1981 Nobel Prize in Physiology or Medicine. "If we had more resources, I would probably fight to see that this activity was not lost." Raising funding for defending human-rights cases has become more difficult than in the past, he adds.

But Quinn believes that the leadership of academic societies too often pitch what he sees as a false dichotomy: that there has to be a choice between human rights and outreach and engagement. Lebowitz agrees, but says

**"One should collaborate, but not overlook abuses."**

that the leaders of the NYAS seem to be honestly convinced that greater engagement and collaborations through outreach may, in the long run, lead to greater improvements in human rights than defending every individual case. "My own feeling is that you should do both," he says. "One should collaborate, but not overlook abuses."

That's a view shared by William Colglazier, who retired a few weeks ago as executive director of the US National Academy of Sciences (NAS), which has an extremely active human-rights committee, and often pursues lengthy behind-the-scenes diplomatic efforts, often at high level within governments. "The NAS's view is that both are extremely important," he says, adding that defending individual cases need not conflict with broader outreach efforts. The NAS is currently working on about 100 human-rights cases. The academy leadership strongly supports such efforts, he says, and they are one of the "most popular activities" with members. He adds that there is "strength" in having a number of organizations supporting any given campaign.

At the NYAS, Wiesel hopes that its outreach programmes could become a useful vehicle for applying pressure in support of individual cases, adding that he is "certainly going to be keeping an eye open to try to see that important cases aren't forgotten".

The NYAS is not alone in reassessing its policy on human-rights issues. In June, the AAAS folded its Science and Human Rights programme into a broader new effort, the Scientific Responsibility, Human Rights and Law Program. This reorganization will help the

AAAS to focus its activities on areas where it can be most effective, such as engaging policy-makers and the public on the social, ethical and legal implications of advances in science and technology, says Jessica Wyndham, associate director of the new programme.

The AAAS had already shut down its substantial programme for taking direct action on cases of persecuted scientists in 2007. Through its new Science and Human Rights Coalition, the association is now educating its members on how to take action on their own to defend individual academics, Wyndham notes.

The AAAS had been one of the few societies, along with the NAS, with the resources to investigate cases in depth. The loss of this capacity is disappointing, says Quinn.

At the ACS, human-rights policy is also in flux. Last year, the ACS's Subcommittee on Scientific Freedom and Human Rights sent five letters defending persecuted scientists to the organization's board for approval. They received no response, says Alec Greer, an organic chemist at Brooklyn College in New York and co-chair of the subcommittee. Last December, the ACS dissolved the committee altogether, and cases will now be reviewed by its Committee on International Activities before the society decides to take action. In a letter sent to the ACS board on 24 May, all ten members of the human-rights subcommittee protested at the dissolution and warned that the new procedures would dilute the ACS's defence of persecuted scientists. Bonnie Charpentier, chair of the ACS, disagreed in a reply dated 5 July: "Far from being dilutive, the new procedures are meant to be more inclusive and effective," she said. The ACS has, since then, sent out one appeal under the new system.

Michele Irwin, administrator of the American Physical Society's Committee on International Freedom of Scientists, says that the human-rights committees of scientific societies bear some responsibility for the changes. She says they have too often worked in isolation from their parent organizations' other programmes, remained overly focused on letter writing, and some have been slow to adapt to a more interconnected world. "I don't think the community has done a good job of convincing people why their work is important, and relevant to what their society is doing," she says. ■



**MORE  
ONLINE**

## TOP STORY



Data from the Tevatron particle accelerator could topple standard model of physics  
[go.nature.com/pzihme](http://go.nature.com/pzihme)

## OTHER NEWS

- Coral genomes could aid reef conservation [go.nature.com/6w4nu6](http://go.nature.com/6w4nu6)
- Bomb-blast brain injuries explained [go.nature.com/35xnj1](http://go.nature.com/35xnj1)
- Alzheimer's disease genes aid the search for preventive drugs [go.nature.com/xxt5gs](http://go.nature.com/xxt5gs)

## Q&A



Iranian AIDS doctor Kamari Alaei still doesn't know why he was imprisoned  
[go.nature.com/nb9ghr](http://go.nature.com/nb9ghr)



## PLANETARY SCIENCE

# NASA picks Mars landing site

*Curiosity rover will explore Gale Crater, which may hold clues to past habitability.*

BY ERIC HAND

**C**uriosity had better strap on its climbing gear. NASA has decided that the US\$2.5-billion Mars rover — the most ambitious mission yet to the red planet — will land in Gale Crater, a 155-kilometre-wide bowl with a mountain at its centre.

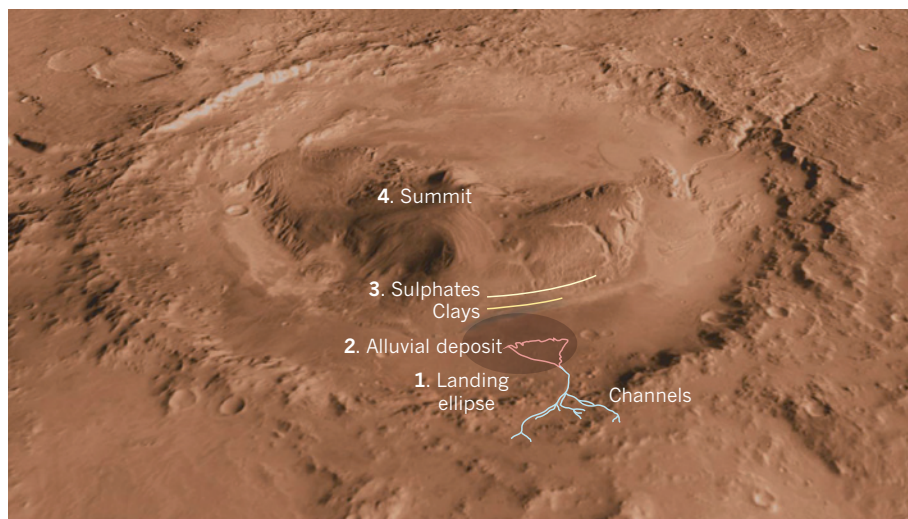
The selection, announced on 22 July, ends a competitive five-year process that considered some 60 sites for the Mars Science Laboratory mission. In the quest to discover whether Mars could ever have supported life, Gale Crater has something for everybody — including a 5-kilometre-tall mountain of stacked sediments.

“This might be the tallest mountain in the Solar System that we can climb with a rover,” says John Grotzinger, the mission’s project scientist and a geologist at the California Institute of Technology in Pasadena.

Four sites, each supported by a vocal group of advocates, reached the shortlist. Gale offered the most to geomorphologists — geologists who study layers of sediment for clues to the processes that laid the layers down. But it also presented ample opportunities to geologists who probe the chemistry of minerals. “Gale is a geomorphologist’s dream,” says Phil Christensen, a Mars researcher at Arizona State University in Tempe. “But it has enough mineralogy to be interesting.”

Curiosity will launch later this year. After a nine-month trip, the 900-kilogram rover, which will carry ten instruments and is built to travel 200 metres a day, will touch down within Gale. Initial plans include a visit to a fan-shaped deposit of sediments that seem to have spilled in through a breach in the crater rim, perhaps transported by water.

The rover will then set out for the mountain. The layered formation was laid down in the crater over several hundred million years, then partially eroded away. Orbiting satellites have spied evidence on the lower slopes of thin layers of sulphates and clays — minerals formed in the presence of water. Determining the abundance and distribution of the minerals, and whether they were formed by groundwater, lake water or precipitation, will help researchers to learn about the habitability of ancient Mars. There are fewer



NASA's Curiosity rover will land (1) in Gale Crater. It will visit an expanse of sediment (2) deposited by channels, then climb a mound bearing water-related minerals (3). It may try to reach the summit (4).

signs of erosion and watery minerals farther up the mountain, but scientists will still want Curiosity to reach the summit. “If you started at the bottom of the Grand Canyon, you wouldn’t stop a third of the way up,” says Christensen. But the climb would take years longer than the mission’s nominal two-year lifetime.

Grotzinger says that all three of the other finalist sites were promising. Eberswalde Crater, which was voted a close second, contains a relic of a river delta and might have offered the best chance of finding preserved organic molecules. Mawrth Vallis, the most ancient of the short-listed sites, was favoured by French researchers who, using an instrument on the European Space Agency’s Mars Express orbiter, found stronger signs of water-formed minerals in the

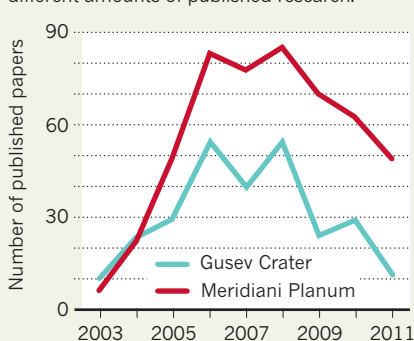
valley than anywhere else on the planet (F. Poulet *et al. Nature* **438**, 623–627; 2005). Finally, Holden Crater, like Gale, may hold ancient lake deposits. “In the end,” says Grotzinger, “we picked the one that felt best.”

Engineers said that Curiosity could be landed safely at any of the four sites, so they left the final choice to the scientists. That placed the burden of choice on Grotzinger and his team — something researchers did not face before the launch of NASA’s Mars Exploration Rovers in 2003. Then, safety considerations ruled out all but two sites — and because there were two rovers, Spirit and Opportunity, both were selected. But the results of that mission demonstrate the importance of site selection (see ‘Landing sites matter’). Opportunity’s observations of Meridiani Planum — a plain rich in the mineral haematite — have produced far more scientific papers than Spirit’s exploration of Gusev Crater. Scientists expected Gusev to hold lake deposits, but it turned out to be covered in comparatively boring volcanic basalt.

Gale Crater was considered for the 2003 rovers. Ultimately, NASA decided that it was too challenging to land a rover in Gale at the time, and the crater was too big for rovers that were expected to traverse only 600 metres over a lifetime of 3 months. But the Opportunity rover has exceeded expectations: it has now entered its seventh year on Mars, and travelled nearly 33 kilometres. If Curiosity has anything like its predecessor’s endurance, the vast and intriguing Gale Crater will probably prove an ample playground for an extended mission. ■

## LANDING SITES MATTER

In 2003, NASA sent two rovers to Mars. Data from the landing sites produced markedly different amounts of published research.

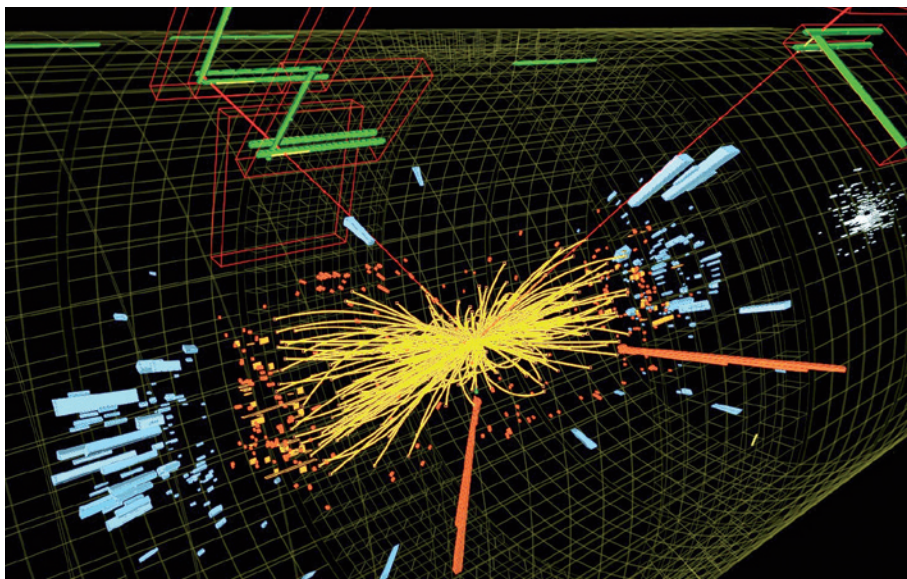


**NATURE.COM**  
Take a virtual tour of  
Curiosity's landing  
site:  
[go.nature.com/3uis7e](http://go.nature.com/3uis7e)



Scan the tag above with the  
free app from **gettag.mobi**





Particle collisions at the Large Hadron Collider — including this smash-up observed by the Compact Muon Solenoid detector — are not yet giving physicists many surprises.

# PARTICLE PHYSICS

# Hint of Higgs, but little more

*No signs of exotic new physics have yet emerged from Europe's giant particle accelerator.*

BY GEOFF BRUMFIEL IN GRENoble, FRANCE

When its experiments started in earnest earlier this year, many scientists hoped that the world's most powerful collider would turn up new particles, additional dimensions and perhaps even a small black hole or two. But beyond a handful of unusual events, the latest data from the Large Hadron Collider (LHC) are frustratingly ordinary.

Based at CERN, Europe's premier high-energy physics lab near Geneva in Switzerland, the LHC accelerates protons to almost the speed of light before slamming them together to create new, heavier particles. For more than a decade, theorists have hoped that the LHC might be powerful enough to generate previously unseen phenomena that would shake the sturdy standard model of particle physics to its core.

But the latest findings from the machine couldn't raise even a tremor inside the main auditorium of the AlpeXpo centre in Grenoble, where scientists gathered last week for the International Europhysics Conference on High Energy Physics. In one session,

Helen Hayward, an experimental physicist at the University of Liverpool, UK, flashed her data from the LHC's ATLAS detector onto the screen, along with the standard model's predictions of the particles that should have emerged from the smash-up. Her observations matched the predictions so perfectly that many of the numbers were identical. "You can see that there's good agreement," she said, with a hint of disappointment. She wasn't alone: in talk after talk, analyses followed the standard model's predictions with unwavering fidelity.

There was one exception. On Friday afternoon, groups working on the two main detectors at the LHC presented evidence of a few extra particles corresponding to something new at energies around 140 gigaelectronvolts (GeV). For now, physicists are only willing to call them "excess events", but if the signal grows stronger as data accumulates, then it could be a sign of the Higgs boson, a vital component of the mechanism that endows other particles with mass. Since the 1960s, scientists have believed that the Higgs, or

**"Supersymmetry is clearly on the ropes."**

something like it, is needed to explain why some particles are heavy and others have no mass at all. The Higgs would also be the key to combining the weak nuclear force — which governs some forms of nuclear decay — with the electromagnetic force, into a single 'electroweak' force. This would see the carriers of the two forces — the W and Z bosons, and the photon, respectively — merge into a single entity at high energies.

But even the Higgs is technically part of the standard model. Instead of confirming the status quo, many physicists anticipated that the LHC might point them in new directions. In particular, theorists hoped the accelerator would turn up evidence supporting a theory called supersymmetry, dubbed SUSY, which postulates a shadowy world of heavy particles corresponding to familiar ones. These superparticles could explain dark matter, mysterious cosmic stuff that seems to interact with the visible world only through gravity. SUSY particles would also eliminate troublesome quantum fluctuations that appear in the standard model and threaten to make nonsense of calculations of the Higgs' mass.

Earlier findings had already cast doubt on SUSY (see *Nature* 471, 13–14; 2011). Now data presented by Hayward and others suggest that superparticles predicted by the most common formulations of SUSY must be heavier than 1,000 GeV. The LHC will probe higher energies as it gathers more data, leaving a chance that it may yet find the superparticles. But even if it did, they would be much too heavy to quell the quantum fluctuations that the theory was originally designed to control. "Supersymmetry is clearly on the ropes," says Rob Roser, a physicist at Fermilab in Batavia, Illinois.

SUSY's supporters say that lighter superparticles could still exist, despite the data, in some formulations of the theory. "I'd say SUSY has a mild hangover for the moment," says Ben Allanach, a theoretical physicist at the University of Cambridge, UK. But, he adds, when the LHC begins operating at its full energy of 14 teraelectronvolts — twice its current level — in a few years time, that hangover could turn into a "fatal migraine".

Others say that data collected during the next six months may be enough to cause serious headaches for theorists. By that point, physicists expect to have gathered enough data to either build up the Higgs signal or prove, once and for all, that the mass-giving boson doesn't exist.

With petabytes of data to be gathered over its 20-year lifetime, the collider could still turn up something entirely new. But many at the meeting admitted disappointment that nothing unusual has popped up inside the machine so far. "I think a lot of people thought there would be some low-hanging fruit," says Roser. ■

➔ **NATURE.COM**  
For a collection of news from the LHC, see:  
[www.nature.com/lhc](http://www.nature.com/lhc)

## GENOMICS

# Genome giant offers data service

*Chinese sequencing institute launches remote computing networks to crunch DNA data.*

BY EWEN CALLAWAY

**T**he world's largest DNA-sequencing outfit is looking to the clouds. The BGI (formerly the Beijing Genomics Institute) this month announced plans to roll out its cloud computing capabilities, which it hopes will help it to dominate bioinformatics in the same way it does the world of sequencing (see *Nature* 464, 22–24; 2010).

Research labs often lack the storage, computing power and technical know-how to cope with the current deluge of genomic data. Services such as the BGI's offer a solution, says Cliff Reid, chief executive of Complete Genomics in Mountain View, California, a competitor of the BGI that sequences and analyses human genomes. "The cloud is going to be central in the entire world of DNA sequencing."

Cloud computing marshals the power of a network of computers that can be accessed remotely to store and analyse data. Creating

a bespoke cloud network to boost the data-crunching power of the Shenzhen-based BGI was a logical move, says Magic Fang, director of the institute's bioinformatics centre. "Competition in sequencing service is becoming more and more intense," he notes. Meanwhile,

***"The cloud is going to be central in the entire world of DNA sequencing."***

the flood of genome data is outstripping scientists' capacity to handle it, as next-generation sequencing techniques drive down the price of reading genomes even faster than the cost of data storage is falling (see 'DNA and chips').

Increasingly, Fang says, scientists who turn to the BGI for its sequencing muscle are also asking for help processing and analysing the data that its next-generation sequencers — which number more than 150 — churn out. By developing its own cloud service, the institute hopes to lure high-profile projects and collaborators,

even if they never set foot in China.

Xueping Quan, a bioinformatician at Imperial College London, made the trip to Shenzhen with raw sequencing data from a plant species whose DNA sequence was too large and complex for her team to assemble into a complete genome. The BGI used its cloud to stitch the data together within a month, and now Quan plans to work with the institute to produce additional genomic data and create a more thorough genome sequence.

At present, the principal use of the BGI's cloud is in the assembly of genomes such as Quan's plant. However, Sifei He, who manages its cloud service, says it will be able to run other bioinformatics software, such as programmes to scour genomes for single-nucleotide variants, or to find places where large chunks of the genome are duplicated or missing.

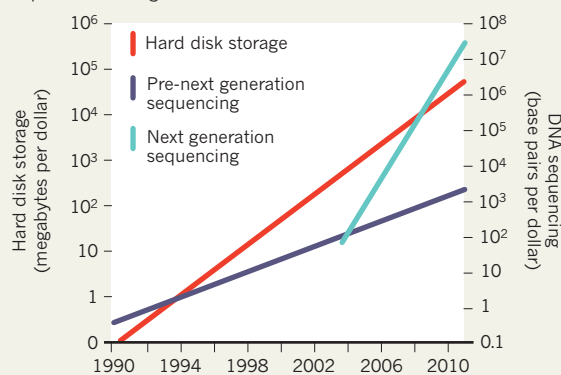
The BGI is not the only sequencing centre pursuing this strategy. Complete Genomics has been storing data for researchers on ►

► Amazon's well established Elastic Compute Cloud for nearly a year, says Reid. He expects that scientists will eventually analyse and store data entirely in such clouds, and only download the final results. Meanwhile, companies such as GenomeQuest of Westborough, Massachusetts — which inked a deal last year to provide bioinformatics to biotechnology giant Syngenta of Basel, Switzerland — and DNAnexus in Palo Alto, California, already run genome-analysis software on clouds.

But it is the BGI's combination of sheer sequencing muscle and in-house cloud computing that makes it stand out as a 'one-stop shop', says David Dooling, a bioinformatician at

## DNA AND CHIPS

The price of DNA sequencing is falling faster than computer storage costs, making cloud computing an increasingly important tool in genomics.



the Genome Institute at Washington University in St Louis, Missouri. "It certainly makes sense," he says. "As more and more people are sequencing and those people are less and less experienced with sequence analysis, you're going to see the proliferation of these vertically integrated solutions."

Paul Flicek, a bioinformatician at the European Bioinformatics Institute near Cambridge, UK, says that it's too early to tell whether the BGI's cloud will address the bioinformatics needs of naive users. But "if somebody can produce a cloud service that's ideal for bioinformatics and costs less than Amazon, there's a niche market there they could really capture", he says. ■

SOURCE: L. D. STEIN GENOMEBIOL. 11, 207 (2010)

## MEASUREMENT

# Physicists count on updated constants

*The latest revision of fundamental quantities bodes well for the proposed overhaul of the international system of units.*

BY EUGENIE SAMUEL REICH

As cheat sheets go, it is a long one. Every four years, a task force of the Committee on Data for Science and Technology (CODATA), headquartered in Paris, releases a short-hand description of the physical world: the latest, painstaking determination of the values and uncertainties of more than 300 physical constants, from the familiar speed of light,  $c$ , to the more obscure mass of the tau lepton,  $\tau$ .

The most recent figures, from 2010, were first placed online in June and announced on 19 July by the US National Institute for Standards and Technology (NIST) in Gaithersburg, Maryland. The numbers include reduced uncertainties for several key constants, which physicists say is encouraging because these will allow for better tests of theory. The more precise figures will also aid plans to redefine familiar units of measurement, such as the kelvin and the kilogram, in terms of unchanging fundamental constants rather than relying on a material object that might not be stable (as for the kilogram standard).

"There's progress in the right direction," says Peter Mohr, a member of the CODATA task

force at NIST. "It gives you more confidence doing science."

The task force revised the numbers that it had released in 2006 by reviewing the latest experiments and theory constraining fundamental quantities. Uncertainties decreased for the fine-structure constant,  $\alpha$ , which is used in atomic physics; Planck's constant,  $h$ , which defines the size of energy quanta; Avogadro's constant,  $N_A$ , the number of atoms in a sample of an element, the mass of which in grams equals its atomic mass; and Boltzmann's constant,  $k$ , which relates energy and temperature.

Much of the progress arises from advances in measurement science. For example, a 2010 measurement of Avogadro's constant made by counting the atoms in a lattice of highly enriched silicon produced a revised number (see *Nature* 467, 892; 2010). This brought the figure closer to another value obtained using Planck's constant (which is related to Avogadro's constant) by measuring the electric current needed to support a kilogram test mass suspended in a magnetic field against the force of gravity. Combining the two methods reduced the relative uncertainty in CODATA's Avogadro's constant from  $\pm 5.0 \times 10^{-8}$  to  $\pm 4.4 \times 10^{-8}$ . For Boltzmann's constant, five new measurements lowered its relative uncertainty to  $\pm 9.1 \times 10^{-7}$  from  $\pm 1.7 \times 10^{-6}$  in 2006, when it was based on just two measurements.

"That's a piece of good news," says Mohr.

The improved precision, particularly in the values of Avogadro's constant and Boltzmann's constant, is a boost for a proposal to rationalize the international system of standard units (known as SI, *Système International d'Unités*) by linking four of them — the ampere, mole, kilogram and kelvin — to physical constants, says Terry Quinn, emeritus director of the International Bureau of Weights and Measures (BIPM) in Sèvres, France. The proposal is due to be voted on at the General Conference on Weights and Measures in Paris in October, and would put the SI on a path to be revamped by 2015.

Plans for the ampere's revision have long been in good shape because measurements of the electron's charge are very precise. Using the latest values for Boltzmann's constant, the kelvin could also be redefined with ease, says Mohr. The sticking point has been the kilogram, which

**"There's progress in the right direction. It gives you more confidence doing science."**

is currently defined by a prototype mass held at the BIPM that many physicists consider an embarrassment because it is thought to vary over time. The new proposal would define the kilogram in terms of Planck's constant, which has units of kilogram metres

squared per second. Before that can happen, some experts would like to see the uncertainty in that constant's value decrease to  $\pm 2 \times 10^{-8}$ .

Quinn says that given recent progress, he expects that to be achieved before 2015. But Ian Mills, a chemist and physicist at the University of Reading, UK, who is president of the BIPM's consultative committee on units, says that the uncertainty of Planck's constant is already low enough to redefine the kilogram reliably. "We could go ahead and make the change to SI today, but everyone is very anxious that we get everything tickety-boo." ■

► **NATURE.COM**  
For more on how to improve the accuracy of SI units:  
[go.nature.com/oujgwf](http://go.nature.com/oujgwf)



# Regulations proposed for animal–human chimaeras

UK lays out first framework to govern ethically sensitive research field.

BY ALISON ABBOTT

The increasingly sophisticated blending of different species to create chimaeras is pushing biology into a new ethical dimension. Last year, scientists used new stem-cell technologies to create a mouse with a functioning pancreas composed entirely of rat cells. So might it soon be possible to create a monkey with a brain composed entirely of human neurons? And would it think like a human?

Such an animal might be useful to researchers studying human cognition or human-specific pathogens. But it would be ethically unacceptable and should be banned, argues a government-commissioned report from the UK Academy of Medical Sciences, a body that promotes medical research.

The document, *Animals Containing Human Material*, says that genetic and stem-cell technologies are now so advanced that the creation of such animals is already on the horizon. But no country has yet devised a broad regulatory framework for the research. The report, released on 22 July, calls for the United Kingdom to take the lead in putting in place specific safeguards.

“We are not proposing a new tier of regulation that will hold up important research,” says Robin Lovell-Badge, a developmental biologist at the Medical Research Council’s National Institute for Medical Research in London, and a member of the working group that drew up the report. At the same time, he says, “we don’t want scientists to cause problems for the future by overstepping the mark of what is publicly acceptable”. Unlike the hypothetical monkey with a human brain, many animals containing human material (ACHMs) are likely to advance basic biology and medicine without transgressing ethical boundaries, the report concludes.

The working group, chaired by human geneticist Martin Bobrow of the University of Cambridge, included experts in philosophy, ethics, social sciences and law, as well as biomedicine, and consulted internationally. The group also commissioned surveys and focus groups that revealed broad acceptance of some mixing of species among lay people who understood the rationale — but also unease about work that could introduce human traits into animals’ brains, reproduction or appearance (see ‘Public support’). The report is likely to inform similar debates in other countries, such as the United

States — which has generated several studies on aspects of ACHM research in the past six years but has no legislation in prospect — and Germany, where bioethical sensitivities are acute.

One category of experiments should be off-limits for the time being, according to the report. This includes the creation of a non-human primate with enough human brain cells to make it capable of ‘human-like’ behaviour. The report says that such animals, which might be able to develop human capacities such as reasoning or self-awareness, would have a moral status close to our own or to that of the

example, enable researchers to introduce a small number of human neural stem cells into a monkey’s brain to assess whether they could replace neurons lost in diseases such as Parkinson’s.

The report urges the UK government to establish a national expert body to advise on ACHMs, and to review the contents of these two categories regularly. The government is likely to incorporate the proposals into legislation.

## INTERNATIONAL ATTENTION

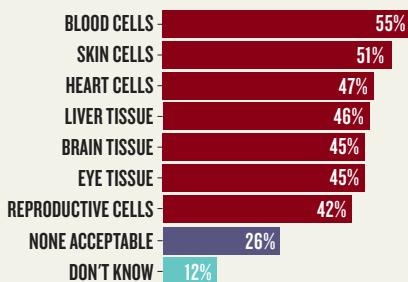
Most ACHM research needs no additional oversight, the report concludes. Innumerable transgenic mice expressing human genes have already been created to study a wide range of human diseases. Transgenic sheep and goats are routinely used to manufacture human proteins for treatment, and pigs containing human genes are being developed for transplantation surgery. In the United Kingdom and most other research-intensive countries, the report says, animal-welfare authorities already regulate this type of work adequately.

Other countries are likely to scrutinize the UK report closely, says molecular biologist Jens Reich, a vice-chairman of the independent German Ethics Council in Berlin, which is preparing its own advisory report on the subject. “ACHM research is very controversial in Germany, and very political because our constitution stresses that the ‘dignity of humans’ must be preserved at all cost,” he says. The UK report concludes that human dignity is not violated by ACHM research.

“It will certainly be looked to in the United States,” says Ruth Faden, a bioethicist at Johns Hopkins School of Public Health in Baltimore, Maryland, who has convened expert groups on the topic. She applauds the UK academy for funding an in-depth public-opinion exercise. “Our debates in the United States would benefit from this type of methodology,” she says. ■ **SEE EDITORIAL P.423 & COMMENT P.448**

## PUBLIC SUPPORT

In a poll of 1,046 Britons last year, about half supported experiments that put human cells or DNA into living animals. There was greatest concern about the use of brain, eye and reproductive cells.



great apes, which cannot be used for invasive research in most countries.

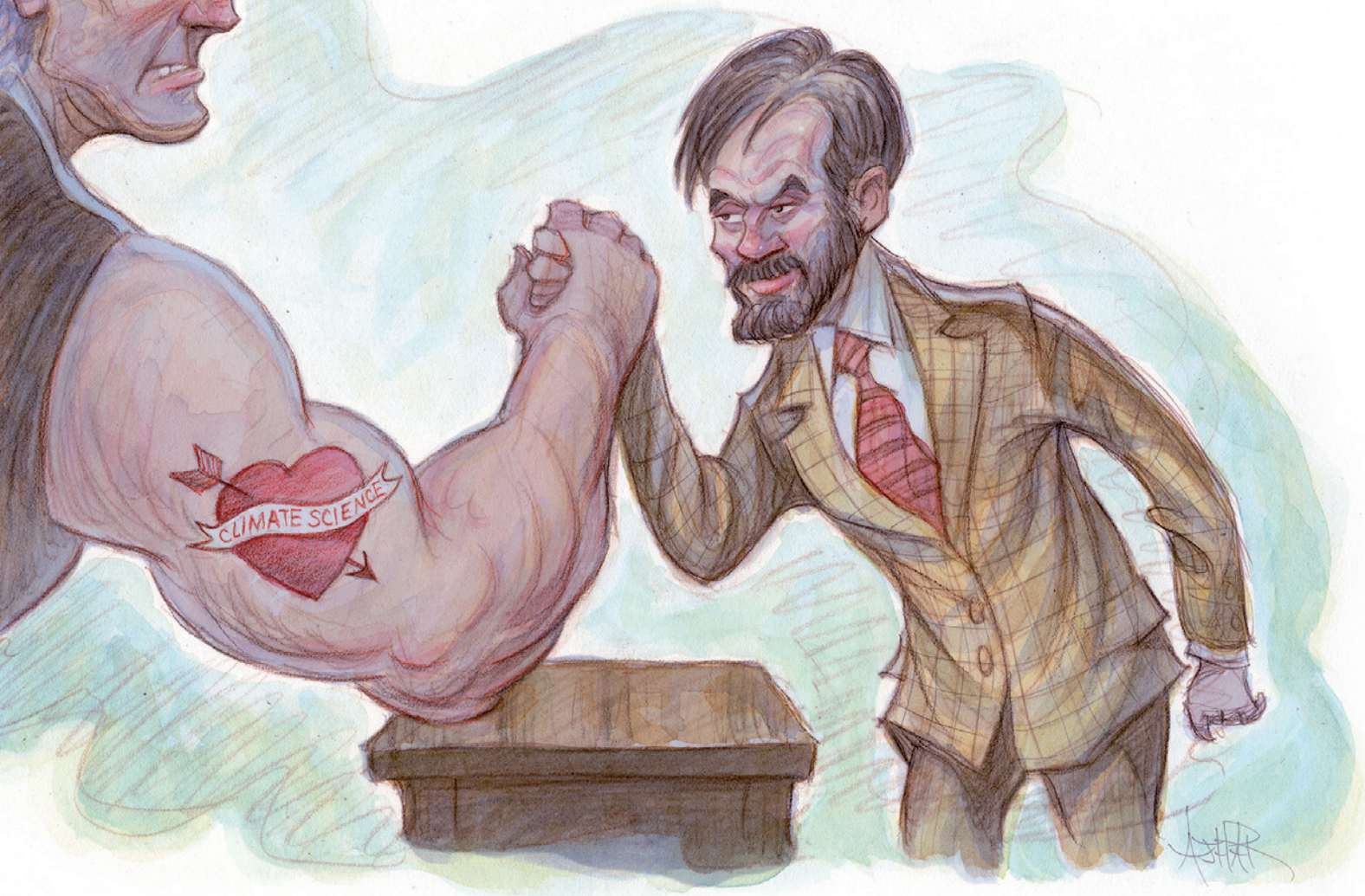
The creation and development of embryos formed by mixing embryonic or pluripotent cells from humans and non-human primates should also be banned for now, as should the breeding of animals that have human-derived sperm or egg cells and could generate a true animal–human hybrid.

A second category of ethically sensitive ACHM research could go ahead if approved by a specialist committee, the report says. This would include modifying the brains of animals, other than non-human primates, in ways that might give some ‘human-like’ function; generating or propagating human-derived sperm or egg cells in an animal in which there is no chance of fertilization; and creating an animal with some obvious human-like characteristic, such as human skin or speech. Some introduction of human genes or cells into non-human primates might also be allowed. That would, for

## CORRECTION

The News Feature ‘A spot of trouble’ (*Nature* **475**, 156–158; 2011) incorrectly states that US federal law requires newborn blood spots to be saved for two years. Only the results of the tests, not the samples, need to be kept on file for that period.





# THE SCEPTIC MEETS HIS MATCH

**Joe Bast and his libertarian think tank are a major force among climate sceptics — but they just can't win the battle over science.**

BY JEFF TOLLEFSON

Joe Bast considers himself an environmentalist. He and his wife Diane used to volunteer for the Sierra Club, one of the United States' largest environmental groups. The two built a geodesic dome in the Wisconsin countryside and once entertained the idea of abandoning the city to live off the land. But today Bast is on the other side of a cultural divide, fighting former colleagues, politicians and scientists on the battlefield of global warming.

As head of the Heartland Institute, a policy think tank with a free-market focus based in Chicago, Illinois, Bast has raised millions of dollars to mount a systematic attack on mainstream climate science. The organization provides fodder for politicians and conservative commentators bent on preventing government regulation of carbon emissions, and Heartland's climate conferences have become rallies for sceptics.

The Heartland Institute, and Bast's personal story, offer a window on the political upheaval taking place in the United States and how it is helping global-warming sceptics to win over some sectors of the public. The shift to the right that enabled conservative Republicans to take over the House of Representatives has killed prospects for comprehensive climate policy in the United States any time soon. Now President Barack Obama's administration is under pressure to scale back a host of regulations on climate, air quality and public health.

All of this contributed to a buoyant mood at Heartland's sixth International Climate Change Conference in Washington DC, which ran from 30 June to 1 July. Speakers presented a litany of accusations against mainstream climate scientists and the Intergovernmental Panel on Climate Change (IPCC). They questioned the validity of many elements of climate science, including the modern temperature

record, palaeoclimate reconstructions and simulations of future conditions.

Since they began in 2008, the conferences have become a prime networking opportunity for those who oppose political action relating to global warming. "For the first time you had this rag-tag army of global-warming sceptics come together to meet each other," says Marc Morano, a former spokesman for Capitol Hill's most renowned sceptic, Senator James Inhofe (Republican, Oklahoma).

At the conference, Bast was upbeat and claimed confidently that publications such as *Nature* and institutions such as the IPCC represent an increasingly entrenched minority. But he nonetheless hesitates when asked about the future. Sitting down for a quiet moment during the conference — because of stress-related health problems, he no longer manages the entire affair — Bast temporarily lets his guard down and acknowledges his

ILLUSTRATION BY BRIAN AHAR



movement's failure to win over mainstream climate scientists. More than ever, they continue to warn about the dangers of human-driven global warming and they retain credibility in the public eye. And Bast points out that the government is still writing cheques that will keep the current research system running.

"We've won the public opinion debate, and we've won the political debate as well," Bast says. "But the scientific debate is a source of enormous frustration."

## PARTY HEARTY

The 2010 election showed the surging appeal of Tea Party activism and its platform of limiting government spending and power, but Bast found his way to a libertarian philosophy long before that. After attending the University of Chicago, he co-founded the Heartland Institute in 1984 at the age of 26. Heartland's annual report says that corporations provided 34% of its US\$6.1-million budget in 2010, with the rest coming from individuals and conservative foundations — some of which have industry ties of their own.

In the past, Heartland has often been criticized for collecting money from tobacco and energy companies, but Bast says Heartland is advocating its own ideology, which generally opposes regulation. He is among the last public defenders of smoking and has argued that concerns about second-hand smoke are as bogus as those surrounding greenhouse gases.

Bast's assault on climate research takes two forms: challenging the credibility of the science, and disputing the claim that there is a scientific consensus on climate change. He does not necessarily deny that humans are having an influence on the climate, but he does question the forecasts of catastrophic impacts and the rationale for curbing carbon emissions.

Heartland plans to spend \$1.8 million on its climate programme this year. Of that, \$413,000 will go to supporting the Nongovernmental International Panel on Climate Change (NIPCC), a small group of sceptics who have set themselves up as a counterweight to the IPCC. Made up of Bast and a few dozen colleagues, the NIPCC mines the scientific literature for nuggets of contrary evidence and doubt — often the kind of uncertainties that scientists readily acknowledge in their publications. The NIPCC also ignores mountains of evidence about the adverse effects of global warming and instead strings together a confident story that makes rising carbon dioxide concentrations seem entirely beneficial.

The group published its first report, *Climate Change Reconsidered*, in 2009, following the release of an executive summary the previous year. Running to more than 800 pages, the report spans the full range of climate science, with a narrative summary of the peer-reviewed literature followed by a detailed list of references. The group plans to release a preliminary draft of its second report next month, and then

a final version in 2012 as a preemptive strike against the IPCC, which will begin rolling out its fifth assessment the following year.

Jay Gulledge, senior scientist at the Pew Center on Global Climate Change in Washington DC, says that Bast, Heartland and the NIPCC all approach scientific data as attorneys, simply trying to sow doubt and justify political inaction. Other climate scientists agree. Gavin

**"The scientific debate is a source of enormous frustration."**

Schmidt of NASA's Goddard Institute for Space Studies in New York says the NIPCC homes in on scientific findings it likes and

then blows them out of proportion.

The last NIPCC report, for example, highlighted a 1999 NASA study<sup>1</sup> that proposed how a negative feedback mechanism over the tropics keeps sea-surface temperatures from exceeding 30°C. The NIPCC extrapolated to say: "If confirmed, this could totally compensate for the warming influence of all anthropogenic CO<sub>2</sub> emissions experienced to date as well as all those that are anticipated to occur in the future." The author of the study, Yogesh Sud of NASA's Goddard Space Flight Center in Greenbelt, Maryland, says that the NIPCC "totally misinterpreted my paper".

Those tactics are not limited to the NIPCC report. Heartland is still circulating a 2007 pamphlet claiming strong disagreement among climate scientists regarding the strength of the scientific case for global warming. The pamphlet cites data from a 2003 online survey<sup>2</sup> by climate scientist Hans von Storch and sociologist Dennis Bray, both affiliated with the Helmholtz Centre for Materials and Coastal Research in Geesthacht, Germany. In the survey, nearly 56% of climate scientists agreed that human activity is causing climate change, 14% were unsure and 30% disagreed. "The survey clearly shows that the debate over why the climate is changing is still underway, with nearly half of the climate scientists disagreeing with what is often claimed to be the 'consensus view'", the pamphlet states.

"We printed 500,000 copies and sent one to every important person in the United States," Bast says, including policy-makers, teachers and business leaders.

But Bast dismisses the findings of a follow-up survey by Bray and von Storch<sup>3</sup>, which found that more than 85% of the responding scientists agreed that human activity is behind climate change.

When the NIPCC released its last report, most scientists paid no attention or rejected it entirely. Michael Mann, a palaeoclimate researcher at Pennsylvania State University in University Park, says the report "is nothing but a mix of myths, half-truths,

cherry-picked distortions, and regurgitated climate-change-denial talking points."

Environmental-policy specialist Roger Pielke Jr at the University of Colorado in Boulder calls it "a big fat bowl of cherries" selectively picked to support the idea that global warming is not a problem. And although Pielke sees the NIPCC as largely irrelevant, he argues that the IPCC has opened the door to such counter-efforts because its most recent assessment did not reflect the extent of the ongoing debates and uncertainties. By failing to be as comprehensive as it could have been, the IPCC "ceded that territory to its critics", he says.

## WAITING FOR THE FALL

Bast happily acknowledges hand-picking data to support his position, but argues that scientists on the other side do the same thing when they are building a case for global warming. He also says it is only natural that a libertarian like him would decide to question the scientific foundation for climate change. Getting serious about global warming means implementing government regulation, going after industry, raising taxes, interfering in markets — all anathema to a conservative agenda. "The left has no reason to look under the hood of global warming," he says. "The right does, and that's what happened."

William O'Keefe, chief executive of the conservative George C. Marshall Institute in Arlington, Virginia, says that more than any other organization, Heartland has kept the focus on the weaknesses in science. But O'Keefe is careful when he talks about what is driving climate politics currently. "I don't think anyone should run to the head of the line in saying they brought the climate legislation to a halt," he says. "I think it was the economy, and when the economy does finally get on a sustainable growth path there will be a willingness to go back and revisit that discussion."

Bast seems to believe that the foundation supporting climate science is collapsing, but with a little prodding, he will talk about his fears, too. Harking back, Bast says he racked up frequent-flier miles throughout the 1990s fending off large government subsidies for sports stadiums all around the country. By the time he made the rounds once, the whole debate would start again. In the end, the stadium projects went forward with taxpayers' money.

"They wore me down," says Bast, "and the same thing may happen with climate change." ■ [SEE EDITORIAL P.423](#)

**Jeff Tollefson** is a correspondent for Nature based in Washington DC.

➔ **NATURE.COM**

For more on climate-change sceptics see: [go.nature.com/a2eebt](http://go.nature.com/a2eebt)

1. Sud, Y. et al. *Geophys. Res. Lett.* **26**, 1019–1022 (1999).
2. Bray, D. & von Storch, H. *The Perspectives of Climate Scientists on Global Climate Change*, GKSS Report (2007).
3. Bray, D. & von Storch, H. *ClSci2008: A Survey of the Perspectives of Climate Scientists Concerning Climate Science and Climate Change*, GKSS Report (2010).



# SETI IS DEAD

---

## LONG LIVE SETI

*The closure of the Allen Telescope Array shifts the search for extraterrestrial intelligence away from big science.*

BY M. MITCHELL WALDROP



ut where the Hat Creek Valley twists among the ancient lava fields north of California's Lassen Peak, the only sounds are the wind and the lowing of distant cattle. The soft growl of antenna motors has long since fallen silent; all 42 radio dishes of the Allen Telescope Array stand motionless in the hot summer sun, staring blindly at the mountains on the southern horizon.

The grass growing around their mounts — its neatness once a point of pride for observatory staff — is getting shaggy. The two caretakers still on site at the Hat Creek Radio Observatory don't have the resources to keep it trimmed. For nearly four years, these dishes listened for radio signals from an alien civilization. But since April, when the state's budget crisis forced the University of California, Berkeley, to suspend operations at the observatory, the world's largest instrument dedicated to the search for extraterrestrial intelligence (SETI) has been left in limbo. If the money cannot be found to reopen the array, the 6-metre antennas will have to be dismantled and removed.

The melancholy vista at Hat Creek makes it easy to entertain equally melancholy thoughts about the SETI enterprise itself. It's the ultimate in high-risk, high-payoff science, pursued by only a handful of passionate researchers. In 50 years of searching, they have turned up nothing — and they can't quite shake an association in the public mind with flying-saucer sightings and Hollywood science fiction, all of which is so easy for cost-cutting politicians to ridicule that any substantial federal funding for SETI is impossible. Private support for the search is getting tighter because of the global recession. And many of the pioneers who have championed the search are now well into their 60s, 70s or 80s.

Given all that, what kind of future can SETI have? Quite a vigorous one, insist the SETI researchers. By nature they are an optimistic lot, given to taking the long view. "I'm a Pollyanna" about raising the money to restart the Allen array, declares Jill Tarter, head of the search programme at the SETI Institute in Mountain View, California — the nonprofit research organization that conceived the array, operated it in partnership with astronomers at Berkeley and is now seeking the funds to resurrect it. Besides, the array is only one SETI initiative among many.

"If it closes I'll be sad for my colleagues at the SETI Institute," says Daniel Werthimer, a radio astronomer at Berkeley who runs several



S. SHOSTAK/SPL

The 42 antennas of the Allen Telescope Array in California listened for alien radio signals for four years, until lack of funding forced the array's closure in April.

SETI surveys using data from the 305-metre radio dish at the Arecibo Observatory in Puerto Rico, “but that’s not going to affect us here.” Or elsewhere: smaller, cheaper SETI searches are currently being conducted on telescopes around the world (see ‘Search engines’). Some are seeking alien radio beacons, whereas others are looking for the flicker of interstellar communication lasers. Some projects are looking at specific stars that seem likely to host Earth-like planets; others are doing a less sensitive but broader scan of the entire sky in the hope of catching signals of a type not yet conceived.

The reason to keep going is in that plethora of projects, says Frank Drake, an early pioneer of SETI studies and now an 81-year-old emeritus professor at the University of California, Santa Cruz. Our Galaxy is vast. The fraction searched so far is tiny. And the conceivable modes of alien communication are myriad. So there are always new possibilities to explore. Or, as Tarter likes to put it, giving up now would be like dipping a cup into the Pacific Ocean, finding nothing but clear water and declaring, ‘the oceans have no fish’.

## INTELLIGENCE TEST

From this wider perspective, the closing of the Allen array would mark the end of a ‘big science’ approach to SETI, but not of SETI itself. And, as the array’s creators ruefully admit, even that approach might have succeeded — if it were not for their own early miscalculations.

Like every modern SETI effort, the Allen array follows the path blazed by Drake in 1960, when he was a staff astronomer at the National Radio Astronomy Observatory in Green Bank, West Virginia, and managed to wangle 200 hours of telescope time to mount the first search for alien signals. Focusing on just two of the closest Sun-like stars, Tau Ceti and Epsilon Eridani, his Project Ozma yielded nothing but static. But Drake’s effort got some prestigious support from physicists Giuseppe Cocconi and Philip Morrison of Cornell University in Ithaca, New York, who had independently come to the same conclusion he had: that massive new radio telescopes like the ones at Green Bank could detect potential alien signals across interstellar distances (G. Cocconi and P. Morrison *Nature* **184**, 844–846; 1959).

Between Drake, Cocconi and Morrison, SETI gained instant scientific credibility, which began to draw other scientists into the field. “I realized I was part of the first generation that didn’t have to ask a priest the ‘Are we alone?’ question,” says Tarter, who committed her career to SETI in 1971 at the age of 27, after reading NASA’s first major report on the subject.

For advocates, the obvious next step was a large-scale, federally funded effort to listen for aliens in as much of the Galaxy as technology allowed. This was a tough sell in Congress, says Drake. “Their first question is, ‘How long will it take?’, and you can’t tell them.” Then they ask how much it will cost — and that depends on how long it takes. “So you can’t guarantee success,” says Drake, “and you’re asking them for a blank cheque.”

But the SETI advocates persevered, and in October 1992 the first phase of NASA’s US\$12-million-per-year SETI search, dubbed the High Resolution Microwave Survey (HRMS), got under way at Arecibo, targeted at the 1,000 nearest Sun-like stars.

A year later it was dead: Senator Richard Bryan (Democrat, Nevada) engineered a bill cancelling the HRMS as an utter waste of taxpayers’ money. “The Great Martian Chase may finally come to an end,” he proclaimed. From then on, says Tarter, “we became the 4-letter ‘S-word’ at NASA headquarters”.

The SETI Institute moved quickly to pick up the pieces. It had been founded in 1984 by scientists from NASA’s Ames Research Center in Moffett Field, California, and had been managing the first phase of the HRMS under contract to Ames. So it had no problem hiring many of the NASA employees who had been working on the programme, and arranging to use the custom-built signal analysers that NASA had already paid for. (These devices were designed to scan radio-telescope data for extremely narrow-band

emissions, which were presumed to be the technological signature of an alien civilization; all known natural radio sources have a much broader bandwidth.)

“We sponsored a big series of workshops between 1997 and 1999 on what the future of SETI ought to be,” says Drake, who served as the first chairman of the SETI Institute’s board of trustees, and is still a board member. One of the prime recommendations was the construction of a large array of small, high-quality radio telescopes that could be dedicated full-time to SETI surveys. As in any radio array, the signals from all the antennas could be combined to look as though they came from a single antenna covering the same area — in this case, spanning 10,000 square metres, or one hectare. The One Hectare Telescope, as it was then known, became the SETI Institute’s top priority.

To help create it, the institute enlisted the telescope-building expertise of the Radio Astronomy Laboratory at Berkeley, where researchers were interested in using the array for conventional surveys of galactic and extragalactic radio sources in parallel with SETI. The partners agreed that the institute would raise the money to build the array, and that Berkeley would design it and pay for 20 years of operation at Hat Creek.

And that, looking back, was when the miscalculations started. “The Allen array was born in a time of irrational exuberance, and ended in the great recession,” says Geoffrey Bower, a radio astronomer at Berkeley who was deeply involved in designing the array. “Those two play a big role in how it was imagined, and now how it’s coming to an end.”

A prime example of over-optimism was the antenna design. The original concept, proposed by Drake, was to use commercial satellite-television dishes as a way to keep the initial antenna cost to an absolute

“YOU CAN’T GUARANTEE SUCCESS, AND YOU’RE ASKING THEM FOR A BLANK CHEQUE.”

minimum. He had been inspired by a 3-metre dish that he had bought and assembled himself for \$600, and that still stands behind his house in the hills near Santa Cruz. “It has receivers on it as good as any radio telescope,” he says. For the array, which was intended to have 350 antennas, Drake found a commercial manufacturer in Wisconsin that would sell him 4.2-metre dishes — good enough, he felt — for about \$1,100 apiece.

But that plan was quickly set aside as the Berkeley team came up with multiple innovations for optimizing performance. Less noise in the receivers. A wider frequency range. More-sensitive measurements of the radio waves’ polarization. More sophisticated electronics.

Tarter and her colleagues at the SETI Institute endorsed these design decisions. But the attempt to do so many new things inevitably led to overruns and delays. The result, despite other innovations to minimize the price tag, was a design for 6-metre antennas that would cost at least \$200,000 apiece. “If we had just reduced the technical complexity of the telescopes early in the project,” says Bower, “we could have built hundreds of them.”

They might have been able to do so anyway — except that the SETI Institute was beginning to realize that it might not have been such a good idea to start the project without getting all the money up front. “We’re scientists,” sighs Tarter. “What did we know?”

The first \$25 million was easy enough to raise through the institute’s contacts in the computer industry. Paul Allen, co-founder of the software powerhouse Microsoft, promised to contribute that much towards design and construction, and the facility was renamed the Allen Telescope Array. But by the time Allen’s first instalment of funds arrived in 2001, the industry was reeling from the collapse of the dot-com boom, making further donations harder to come by. And those computer-industry philanthropists who were still giving, preferred to fund projects

➔ **NATURE.COM**  
To listen to the  
author discuss SETI  
in a podcast, visit:  
[go.nature.com/apmibm](http://go.nature.com/apmibm)



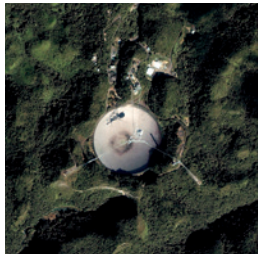
Scan the tag above with the  
free app from [gettag.mobi](http://gettag.mobi)



## SEARCH ENGINES

With the US\$50-million Allen Telescope Array shuttered, the search for extraterrestrial intelligence (SETI) is now sustained by 'small-science' efforts that use borrowed telescope time and inexpensive detectors to seek radio and optical signals from aliens.

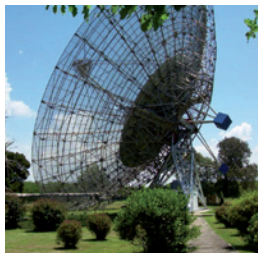
### RADIO SETI



#### SETI AT BERKELEY

Analyses data from the Arecibo radio telescope in Puerto Rico and includes SETI@home, which uses processing power donated by home computers.

**Cost:** ~\$175,000 plus ~\$550,000/year  
**Funder:** Various, including SETI@home participants



#### SOUTHERN SETI

One of the few Southern Hemisphere SETI projects, this uses two 30-metre antennas at the Argentinian Institute of Radio Astronomy, near Buenos Aires.

**Cost:** ~\$16,500/year  
**Funder:** The Planetary Society in Pasadena, California

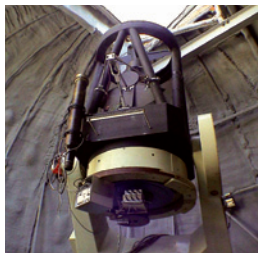


#### PROJECT ARGUS

Coordinating SETI observations by hobbyist radio astronomers, this project currently comprises 147 home-built dishes.

**Cost:** ~\$4,000/dish  
**Funder:** Participants

### OPTICAL SETI:



#### SETI AT BERKELEY

A detector on a 0.76-metre telescope at the University of California, Berkeley's Leuschner Observatory searches for nanosecond pulses of laser light.

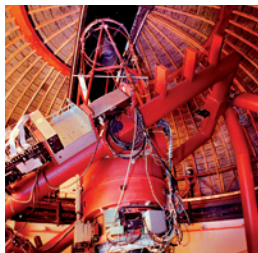
**Cost:** ~\$70,000 plus ~\$12,000/year  
**Funder:** The SETI Institute in Mountain View, California, and Planetary Society



#### SETI OPTICAL TELESCOPE

This 1.8-metre telescope in Harvard, Massachusetts, also looks for nanosecond laser pulses.

**Cost:** ~\$350,000 plus ~\$20,000/year  
**Funder:** SETI Institute; Planetary Society; and Bosack-Kruger Charitable Foundation in Redmond, Washington



#### LICK OBSERVATORY OPTICAL SETI

A detector on a 1-metre telescope at Berkeley's Lick Observatory looked for laser pulses in 2000–07. Berkeley is seeking \$100,000 to resume the search.

**Cost:** ~\$20,000 plus ~\$12,000/year  
**Funder:** SETI Institute and Planetary Society

with a nearer-term, more tangible payoff. "We don't have patients we can cure," says Thomas Pierson, the SETI Institute's chief executive, summing up the problem.

In the end, the institute was able to raise only about \$50 million in total, roughly \$60 million shy of what it would need to pay for all 350 dishes. The array began operations in 2007 with just 42 antennas. This lack of collecting area greatly slowed down the alien-hunting. And worse, it left the array without the sensitivity required for cutting-edge radio astronomy — a fact cited by the National Science Foundation in 2008 when it declined to renew the grant that the university had been using to fund Hat Creek operations. Suddenly, Berkeley had no way to pay the array's \$2.5-million annual operating costs.

The partnership managed to limp along for a few more years. But its increasingly urgent efforts to find a permanent source of funding ran headlong into the worldwide economic downturn and California's resulting budget crisis. On 22 April 2011, Berkeley and the SETI Institute were forced to pull the plug on Hat Creek.

### THE SEARCH CONTINUES

Berkeley doesn't hold out much hope of reopening the array, and has started to look at other radio-astronomy projects. "We've poured everything we had into the array," says Bower, "and really hit a wall with it."

The SETI Institute, though, is still trying. In June, it launched an experiment in 'crowd-sourced' funding on a website called SETIstars.org, through which individual supporters can make donations of \$10–1,000. The hope is to raise a few hundred thousand to a million dollars per year that way (the take so far is just over \$100,000). The institute is also in ongoing talks with the US Air Force, which is interested in providing some operations money in return for part-time use of the array to track debris in orbit around Earth.

But if the closure proves to be permanent, say institute officials, their plan B is to fall back on smaller-scale efforts — in effect, turning to the methods that the rest of the SETI community has followed all along. This approach tends to be very ad hoc and informal, says Werthimer. It's done with limited funding, borrowed telescope time and investigators who, like Werthimer himself, do SETI only part-time. "And that's the way it should be," says Werthimer. "It's naive to think that we know what ET, a billion years ahead of us, is going to be doing. So we want to be a small-scale science, trying lots of things."

Data from the Arecibo telescope, for example, are sent to five receivers at once: three doing conventional radio surveys, and two, operated by Werthimer's group, doing SETI. At the Oak Ridge Observatory in Harvard, Massachusetts, Paul Horowitz — a Harvard University astronomer — and his students are searching for extraterrestrial laser pulses with a small optical telescope, built for the search with roughly \$300,000 from nonprofit organizations.

"I see SETI as a terrific thing for a graduate student," says Horowitz. "It's an unploughed field." Students can think up a plausible mode of extraterrestrial communication; design, build and test a detector; analyse the data; and get tangible, hands-on experience. Most students in conventional astronomy just get observation time on a big, institutional telescope. "When I first heard what Dan was doing — wow! I couldn't imagine doing anything else," says Andrew Siemion, who is writing his PhD thesis on SETI instrumentation under Werthimer.

Perhaps this is the lesson from the mothballed array at Hat Creek: the odds against success with SETI are so long that it is best done as a small-science, part-time pursuit. Or perhaps not: the SETI Institute may yet find a way around all the funding woes and political headwinds, and bring the Allen array back to life.

Either way, SETI is not going to disappear. "I went into SETI because it was just too cool not to," says Horowitz. And others will too, he says. "All you need is a university environment, with a few tenured faculty willing to host something wacky. The question is always interesting, and there will always be people willing to take a long shot." ■

**M. Mitchell Waldrop** is a features editor for *Nature*.

# COMMENT

**POLICY** Research with animals containing human material needs regulation **p.448**

**CONSERVATION** Declining shark numbers implicate fin hunters **p.451**

**CULTURE** Contemporary artists inspired by psychiatric patients **p.452**

**FISHERIES** European reforms should follow the money and scientific advice **p.454**



TORU HANA/REUTERS/COREIS



Carbon markets tackle factory emissions in many nations around the world; the European Union is now looking to include non-EU aviation.

## Carbon trading needs a multi-level approach

International agreements are not enough. Interlinked national and regional tools are also needed to reduce carbon emissions, say **Mark Maslin** and **Joanne Scott**.

Carbon emissions rose to an all-time high in 2010, and international negotiations to agree binding national carbon-emissions targets have so far failed. However, intergovernmental failings have not slowed the expansion of trading in carbon emissions.

Carbon markets are being set up at multiple levels of governance — including locally, regionally and supranationally. Markets have been set up in the European Union (EU) and New Zealand, and plans are afoot for markets in California, Australia

and India. The biggest by far, the European Union's Emissions Trading Scheme (ETS), is using its carbon market power to experiment with new solutions to old problems, and to spur climate action elsewhere.

Controversially, the ETS is set to extend its carbon market to include some non-EU emissions. For example, from 2012, the ETS could cover emissions amounting to almost 60% of the world's international commercial aviation emissions. Although this move is being challenged in the courts, the EU claims that it is already encouraging the

United Nations International Civil Aviation Organization to consider a global regime for aviation emissions. Similarly, the EU is using the threat of further ETS expansion to galvanize global carbon regulation of maritime transport.

Such overlapping, interacting regulatory frameworks have many advantages. They offer multiple opportunities to learn which regulations are most effective in the short and long term. Local efforts also build redundancy into the system: they act as regulatory safety nets if higher-level ►



► negotiations break down. And, as shown by the EU, local action can help to shape the dynamics of international negotiations. Critics fear multiple bureaucracies that are messy and confusing, and which introduce high transaction costs, but although a top-down system of governance may be neat, such an effort would be unlikely to succeed.

### THE RISE OF INTELLIGENCE

Trading in carbon emissions has created a major global market since the mid 1990s, with an estimated spend of US\$142 billion in 2010, of which 85% was through the ETS<sup>1</sup>. Another \$48 billion was spent in 2009–10<sup>2</sup> solely on carbon-market intelligence (gathering and analysing information required for the effective sale or purchase of carbon-based credits). The size of the carbon-emissions market itself was essentially unchanged between 2009 and 2010, despite the economic recession (see ‘The size of the carbon market’). At the same time, businesses and governments have been increasingly investing in efforts to understand the flow of carbon, in order to plan economic and environmental strategies. Updated figures compiled for this article by UK-based consultancy kMatrix show that the intelligence sector has grown again in 2010–11, to \$56 billion, a growth of 16% in a single year (see ‘The size of the carbon market’). The quality and usefulness of this information varies enormously, but its growth underscores the importance of reliable carbon data in understanding the development of this potentially huge market.

The bulk of the carbon-emissions market is made up of trading in carbon allowances in ‘cap-and-trade’ systems. These markets are usually national and regional because it is easier to set a cap on emissions within borders and in relation to particular industries. At the international level, carbon markets trade in credits from certified projects under the Kyoto Protocol’s Clean Development Mechanism (CDM). The CDM allows investors, such as companies and countries, to purchase carbon offsets known as Certified Emission Reductions by investing in carbon-reduction projects in developing countries. For example, the carbon credits from the emissions saved as a result of the construction of a wind farm in India can be bought by an EU power company to meet its emissions-reduction obligations under the ETS.

Assessing — monitoring, recording and verifying — the quality and quantity of carbon emissions ‘saved’ has long been a problem for the CDM. For example, credits for projects involving the capture of industrial gases (hydrofluorocarbons or HFCs) have been regrettably easy to game. The regulation has created a perverse incentive for companies to produce more HCFC-22, a refrigerant and powerful greenhouse gas being phased

out under the Montreal Protocol, in return for windfall profits for capturing the HFC-23 by-product from its production<sup>3</sup>. About 70% of Certified Emission Reductions have come from projects of this kind. Depressingly, the European Commission concluded this year that production of HCFC-22 could be higher today than it would have been in the absence of CDM activity.

The problem with industrial gas credits has been known since at least 2007, yet the multinational executive board that supervises the CDM has been unable to reach an agreement on how to solve it. So far, only a partial and provisional ban on gas-capture projects has been agreed. Happily, the EU does not rely exclusively on international regulation of the CDM. Having finally lost patience with international regulators, it will ban the use of credits from industrial gas projects in the ETS from 2013.

The industrial gases saga has highlighted other failings of the CDM. Perhaps the hardest problem for carbon offsets concerns ‘additionality’. In principle, Certified Emission Reductions come from projects that create emissions reductions that are ‘additional’ to those that would otherwise have been achieved. Verifying additionality by reference to a baseline, which rests on a business-as-usual assessment (the level of carbon

*“The EU has lost patience with international regulators.”*

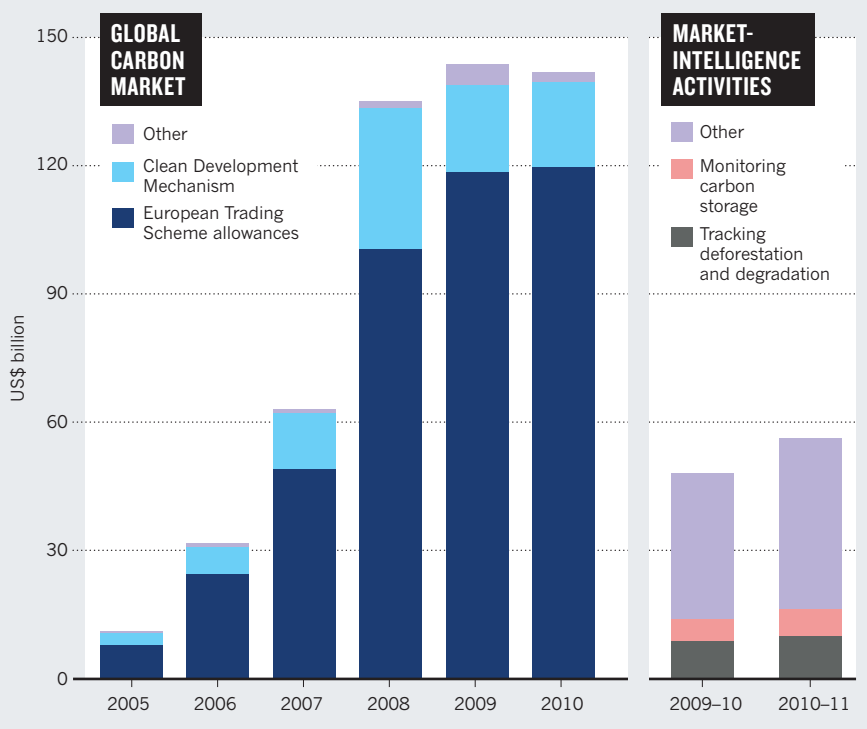
emissions if this project had not been put in place) has proved very difficult. As a result, there are doubts about the CDM’s ability to bring about genuine emissions reductions.

One solution to the additionality problem may come from the EU push into sectoral carbon trading. A ‘sectoral’ approach to emissions would fundamentally change the way in which the EU measures emissions saved as a result of CDM projects. Additionality would no longer be assessed by reference to a hypothetical business-as-usual baseline, but by reference to sector-by-sector carbon-efficiency standards. To generate credits, a project would have to cut emissions below the level of ‘mere’ compliance with the standard. In the absence of global agreement, the EU is seeking bilateral agreements to move forward with this approach, and is threatening to refuse Certified Emission Reductions from new projects after 2013 — except those based in least-developed countries — to persuade countries to come on board.

The EU is also extending the geographical reach of the ETS. From 2012, all flights arriving in or departing from an EU airport will be required to surrender an emissions allowance or certificate for each tonne of carbon emitted during flight, including those occurring outside EU airspace (unless those emissions are subject to climate regulation elsewhere). By 2020, the ETS could add an estimated \$11–56 to a return long-haul flight. At first, 85% of these certificates will

### THE SIZE OF THE CARBON MARKET

Carbon-emissions trading was growing fast until it stalled when global recession hit in late 2008. But spend on market-intelligence activities, from project development to carbon accounting, grew by 16% between 2009–10 and 2010–11.



SOURCES: REF. 1: KMATRIX

be issued by the EU free of charge, with the airlines buying the remaining 15% at auction. These rules are being challenged by the major US airlines in a case currently before the EU's highest court. A final judgment is expected in the first half of 2012.

Policy-makers need examples of good practice or good ideas. The International Carbon Action Partnership — a group of countries and regions that are pursuing cap-and-trade systems — is intended to share knowledge and experience on the development of carbon markets. Similarly, Sandbag, a London-based non-governmental organization, submitted comments in 2010 on California's plans to establish a cap-and-trade system, which drew heavily on the good, the bad and the ugly of European experience with the ETS.

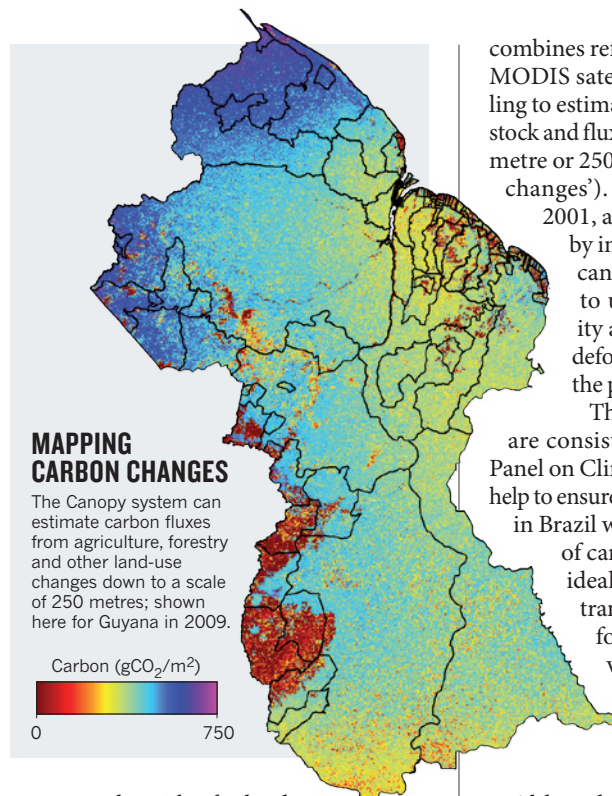
### IF NO ONE HEARS A TREE FALL

Lessons from the ETS experience with the CDM should be applied to the design and operation of the UN programme for Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (REDD+). Although there is debate about the share of global carbon emissions that may be attributed to deforestation and land-use change, a figure of around 20% is often put forward<sup>4</sup>. Without proper oversight at multiple levels of governance, as with the CDM there will be real difficulties in assessing the quantity and quality of the carbon emissions 'saved'.

Following the UN climate negotiations in Copenhagen in 2010, developed countries made non-binding pledges to reduce emissions by between 13% and 19% (compared with a 1990 baseline)<sup>5</sup>. The EU has made a binding commitment to reduce emissions by 20% by 2020 (again compared with a 1990 baseline). If carbon allowances or offsets from REDD+ were available for use in achieving compliance with these targets, the need, worldwide, for meaningful domestic cuts would cease. Hence, it is crucial that there are robust ceilings on the use of REDD+ credits in domestic carbon markets, if incentives for national action are to remain meaningful.

It is also essential that the projects funded do not generate adverse social, cultural or environmental effects. Investors in REDD+ and governments that permit REDD+ allowances to be used for compliance must share responsibility with host governments for developing and improving project assessments to avoid negative effects. There are also worries that REDD+ underestimates implementation costs and could undermine good forest governance programmes in key countries<sup>6</sup>.

Again, the biggest and most intractable problem is additionality. Although 47% of anthropogenic carbon emissions remain in the atmosphere, 27% are absorbed by the



oceans and 26% by the land biosphere<sup>7</sup>. For a REDD+ carbon credit, something 'additional' must be done to prevent future loss of land carbon stock, for example, through avoided deforestation. Countries that take early and effective action to preserve their biosphere boost their baseline efforts, perversely making it harder to get financial rewards from REDD+. Current UN negotiations are considering possible ways of addressing this problem, including setting a regional deforestation rate, which would serve as a baseline for all countries in a given region, regardless of their actual or projected rate of deforestation.

The multiple challenges of REDD+ will not be solved by the UN alone. Until recently, measuring the change in land carbon stock and annual flux was thought to be 'beyond science', because only static data sets for the early 2000s were available<sup>4</sup>. Despite this, a staggering \$6.3 billion was spent in 2010–11 (a 20% increase on the previous year) on monitoring land carbon, with the quality of data often being extremely poor. The lack of verifiable data for REDD+ has the potential to undermine the whole programme by making it easy to cheat. However, considerable scientific progress is being made and new methods for verifying data are emerging.

One of us (M.M.) is involved in such an approach. Work by Carbon Auditors, a London-based company spun out from University College London, in partnership with business service company Logica, shows that monitoring, recording and verification can be done effectively using their jointly developed 'Canopy' system. Canopy

combines reflectance data from the NASA MODIS satellite with ecosystems modelling to estimate monthly-to-annual carbon stock and flux at a global resolution of 1 kilometre or 250 metres (see 'Mapping carbon changes'). Because the data go back to 2001, and accuracy can be enhanced by including ground data, Canopy can be used to draw up baselines, to understand natural variability and to identify major areas of deforestation and degradation over the past decade.

The development of systems that are consistent with Intergovernmental Panel on Climate Change guidelines<sup>8,9</sup> will help to ensure that one tonne of carbon saved in Brazil will be equivalent to one tonne of carbon saved in Australia. In an ideal world, there will be multiple, transparent, overlapping schemes for monitoring, recording and verification, with domestic and third-party organizations contributing to a sense of confidence in the emissions saved and the credits accrued.

Although we may continue to pin our hopes on the emergence of a comprehensive global agreement on climate change, we should recognise that even if this were to be achieved, multi-level governance would still be required. As emissions trading markets grow, governments at all levels must take responsibility for ensuring that such trading is delivering properly on its goals. ■

**Mark Maslin** is in the Department of Geography, University College London, Pearson Building, Gower Street, London WC1E 6BT, UK. **Joanne Scott** is in the Faculty of Laws, University College London, Bentham House, Endsleigh Gardens, London WC1H 0EG, UK.  
e-mail: mmaslin@geog.ucl.ac.uk

1. Linacre, N. et al. *State and Trends of the Carbon Market 2011* (World Bank, 2011); available at: [go.nature.com/xveyp4](http://go.nature.com/xveyp4)
2. Poessinouw, M. & Howard, S. *Knowledge Matrix Limited Carbon Market Intelligence Report for Carbon Auditors Limited* (2010).
3. Wara, M. *Nature* **445**, 595–596 (2007).
4. Saatchi, S. S. et al. *Proc. Natl Acad. Sci. USA* **108**, 9899–9904 (2011).
5. Levin, K. & Bradley, R. *Comparability of Annex I Emission Reduction Pledges*. Working Paper (World Resources Institute, 2010); available at: [go.nature.com/m7uipe](http://go.nature.com/m7uipe)
6. Phelps, J., Webb, E. L. & Agrawal, A. *Science* **328**, 312–313 (2010).
7. Friedlingstein, P. et al. *Nature Geosci.* **3**, 811–812 (2010).
8. Penman, J. et al. (eds) *Good Practice Guidance for Land Use, Land-Use Change and Forestry* (IPCC, 2003).
9. Eggleston, S. et al. (eds) *2006 IPCC Guidelines for National Greenhouse Gas Inventories. Volume 4 — Agriculture, Forestry and Other Land Use* (IPCC, 2006).

**The authors declare competing financial interests:** for details see [go.nature.com/h8k4dk](http://go.nature.com/h8k4dk)



# Regulate research at the animal–human interface

The time is right, says **Martin Bobrow**, to improve the governance of research involving animals that contain human genetic or cellular material.

Chinese scientists have introduced human stem cells into goat fetuses, producing animals with organs that contain functioning human-derived cells<sup>1</sup>. US scientists have examined the ethics of creating a mouse that has some human-derived brain cells (although they have not done the experiment)<sup>2</sup>. Many thousands of transgenic rodents, inter-species cell lines and animals grafted with human tissue have been created worldwide, and have helped substantially with the investigation of health and disease and the development and testing of therapies.

But few countries have specifically considered the governance of research involving animals that contain human material (ACHM), and the topic has had little public discussion. Potentially controversial science proceeds best in an open environment. Britain has an enviable record of reasonably harmonious research in human embryology owing to parliamentary debate, legislation and regulators who systematically engage the public. However, violent opposition to animal research has sometimes hampered open discussion. We hope that era is behind us, and the opportunity for inclusive discussion of these more subtle issues can be grasped.

And it must be grasped, because in Britain at least, there is a regulatory discontinuity. Research with embryos that contain animal material but are judged by regulators to be 'predominantly human', is subject to stringent scrutiny and authorization by the Human Fertilisation and Embryology Authority (HFEA). The HFEA takes account of scientific, medical, ethical and social issues, and frequently consults the public on emerging techniques. A similar chimaeric embryo with marginally less human material judged 'predominantly animal', however, is regulated by the Home Office in consultation with the Animal Procedures Committee under legislation intended to protect animal welfare. In mammals, this regulation becomes applicable at the midpoint of gestation.

The lack of a formal working interface between these two systems creates uncertainty for scientists requiring regulatory approval for their work. It could also allow sensitive experiments to be done legally but without expert ethical scrutiny.

To help bridge this regulatory boundary,



Enzymes in the milk of this genetically altered goat will combat diarrhoea-causing bacteria.

on 22 July, the Academy of Medical Sciences in London published a report<sup>3</sup> of a working group that I chaired. The report recommends where the ethical limits of such methods may lie, and what governance is needed. We hope that it will catalyse the development of international standards and guidelines by collaboration among regulators, policy-makers, bioethics bodies, funders of medical research and the research community.

## SENSITIVE AREAS

To inform the report, the academy commissioned a public dialogue that involved more than a thousand people<sup>4</sup>. Most participants supported ACHM research that is aimed at better understanding or treating human disease. But three areas of sensitivity emerged: modifications of the animal brain that are likely to lead to human-like cerebral function; experiments that could lead to functional human gametes in an animal (especially if the gametes might be fertilized); and modifications to an animal that create features perceived as uniquely human, such as facial shape, skin texture or speech.

For ACHM, the academy recommends a similar regulatory structure to that proposed internationally for human stem-cell research. First, the great majority of experiments are uncontroversial and should undergo standard ethical and regulatory review. Second, proposed experiments that approach these areas of sensitivity should be scrutinized by a national expert multidisciplinary body that also advises on more general aspects

of animal research. Such sensitive research should proceed in incremental steps, with regular dialogue between researchers and regulators. Third, a very limited number of studies should not currently be undertaken because they raise very strong ethical concerns or lack sufficient scientific justification.

UK research regulations must be revised by 2012 to transpose the European Directive on the protection of animals used for scientific purposes into UK law. At the same time, the UK government is reducing the number of its public agencies (including, perhaps, the HFEA). We hope the Home Office will take this opportunity to reshape the Animal Procedures Committee into a body with a remit to provide guidance on these ethically and socially sensitive kinds of research. Recognizing that an effective regulatory system must not needlessly hamper potentially beneficial science, we recommend that this body should be sufficiently flexible and consultative to adapt to evolving scientific knowledge and social attitudes.

Securing a robust, forward-looking regulatory framework for ACHM would promote Britain's position as a responsible home for cutting-edge science. But science is an international endeavour. Other countries must follow suit. Consistency in regulation and governance promotes constructive multinational collaboration and ethically responsible research, as the human-stem-cell research community is demonstrating through the production of common guidelines. The harmonization of animal-research regulations across Europe provides a stimulus to begin a similar process for ACHM. ■

**Martin Bobrow** is *emeritus professor of medical genetics at the University of Cambridge, UK, and chair of the Academy of Medical Sciences working group.*  
e-mail: [Laura.Boothman@acmedsci.ac.uk](mailto:Laura.Boothman@acmedsci.ac.uk).

1. Zeng, F. *et al.* *Proc. Natl Acad. Sci. USA* **103**, 7801–7806 (2006).
2. Greely, H. T., Cho, M. K., Hogle, L. F. & Satz, D. M. *Am. J. Bioeth.* **7**, 27–40 (2007).
3. Academy of Medical Sciences *Animals Containing Human Material* (2011); available at [go.nature.com/uy17hx](http://go.nature.com/uy17hx)
4. Ipsos MORI *Exploring the Boundaries: Report on a Public Dialogue into Animals Containing Human Material* (2010); available at [go.nature.com/oucud](http://go.nature.com/oucud)

AP PHOTO/R. PEDRONCELLI



BETTMANN/CORBIS

A US hydrogen bomb lost in the sea off Spain in 1966 sparked a search using Bayes' theorem. It was eventually recovered after a fisherman tipped off authorities.

## STATISTICS

# Known unknowns

Andrew Robinson enjoys a history of a controversial probability tool — Bayes' theorem.

Bayes' theorem of probability was proposed by English mathematician and clergyman Thomas Bayes in the 1740s, and rediscovered in the 1770s by Pierre Simon Laplace, a French mathematician. It states that by updating our initial beliefs with new objective information, we get an improved belief. Or as economist John Maynard Keynes put it: "When the facts change, I change my opinion."

Considering the widespread effectiveness of Bayesian inference in physics and astronomy, genetics, imaging and robotics, Internet communication, finance and commerce, it is surprising that it has remained controversial for so long. Many twentieth-century scientists who used the Bayesian approach in their work — including mathematician Alan Turing and physicists Enrico Fermi and Richard Feynman — declined to use the 'B' word in public.

Sharon Bertsch McGrayne explains their reticence in her impressively researched history of Bayes' theorem, *The Theory That Would Not Die*. The statistical method

runs counter to the conviction that science requires objectivity and precision, she writes. Bayes' theorem "is a measure of belief. And it says that we can learn even from missing and inadequate data, from approximations, and from ignorance."

A crucial example of the application of the theorem was Turing's cracking of the German

**"High-speed computing is the main reason that Bayesian methods shook off their detractors."**

naval cipher Enigma during the Second World War, which played a key part in the Allied victory in 1945. After the war, Turing's wartime assistant, I. J. 'Jack' Good, wrote about Turing's Bayesian technique for finding pairs and triplets of letters in the cipher. To avoid censorship under the UK Official Secrets Act, he described it in terms of bird watching.

Suppose a birder spotted 180 different species, many of which were represented by only one bird. Logically, other species must have been missed. A frequentist statistician

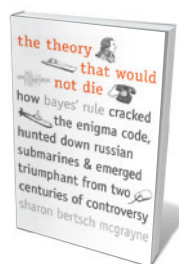
would count those unseen species as zero, as if they could never be found. Turing, by contrast, assigned them a tiny non-zero probability, thereby factoring in that rare letter groupings might not be present in his current collection of intercepted messages but could appear in a larger sample. The same technique was later adopted in DNA sequencing and by artificial-intelligence analysts.

Turing never mentioned Bayes; he may not even have been aware of the theorem. McGrayne speculates that Turing either rediscovered the idea himself in 1940, or heard about it through geophysicist Harold Jeffreys at the University of Cambridge, UK, who published his Bayesian *Theory of Probability* in 1939. Good told McGrayne in an interview before his death in 2009 that he had asked Turing whether he was essentially using Bayes' theorem. To which Turing apparently replied: "I suppose."

The advent of the cold war put Bayes' theorem into cold

**NATURE.COM**  
How Bayes' theorem is used in court:  
[go.nature.com/bcu1h6](http://go.nature.com/bcu1h6)





**The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines and Emerged Triumphant from Two Centuries of Controversy**

SHARON BERTSCH  
MCGRAYNE  
Yale University Press:  
2011. 336 pp. \$27.50

storage. Publication of the method used to decode Enigma was embargoed under the Secrets Act until the 1990s. During the 1950s, the influential statistician and geneticist R. A. Fisher continued his long-standing battle against Bayes' theorem, calling it an "impenetrable jungle". When Good discussed the theorem at Britain's Royal Statistical Society, the next speaker began: "After that nonsense..."

In the United States, the small group of Bayesian statisticians came under suspicion as outsiders. During the McCarthy period

of anti-communist sentiment, they were even considered 'un-American'. Professors at Harvard Business School referred to their Bayesian colleagues as "socialists and so-called scientists".

Meanwhile, the US military and government were both applying Bayes' theorem, if reluctantly. In a gripping chapter, McGrayne describes how it was used in trying to find a hydrogen bomb that fell from a B-52 jet into the sea off Spain in 1966, and a US nuclear-powered attack submarine that disappeared in the Atlantic Ocean in 1968. President Lyndon B. Johnson raged to the investigating team: "I don't want this probability stuff. I want a plan that tells me exactly when we're going to find this bomb."

In the event, Bayes' theorem was not needed. The missing bomb was located in an ocean canyon following a tip from a fisherman who had seen a parachute splash down near his boat. As for the Bayesian search for the submarine in the Atlantic, the consensus is that it would have succeeded had faster computers been available in 1968. High-speed computing is the main reason that Bayesian methods shook off their detractors and acquired their present-day prominence, McGrayne emphasizes.

For all the book's skilful mingling of ideas and intriguing personal details, I found it sloppy on occasions. Tautologies slip in, and evidence is lacking for some claims. Nonetheless, *The Theory That Would Not Die* is a rollicking tale of the triumph of a powerful mathematical tool. ■

**Andrew Robinson** is author of *The Man Who Deciphered Linear B: The Story of Michael Ventris*.  
e-mail: andrew.robinson33@virgin.net



IMAGESOURCE/PHOTOLIBRARY

Sharks are more threatened by humans than we are by them, yet the great white still gets a bad press.

MARINE SCIENCE

## Saving the shark

**Steven Campana** enjoys a refreshing account of human pressure on the oceans' top predator.

**H**ow do shark-attack deaths differ from those caused by chairs, toasters and falling coconuts? As science journalist Juliet Eilperin explains in *Demon Fish*, you are about 200 times less likely to be killed by a shark than by a defective toaster.

Her book delves into the uneasy cohabitation of two of Earth's top predators: sharks and humans. People used to coexist with sharks more amiably before the advent of technology, even deeming them gods. But with films such as *Jaws* as a guide, many now presume that sharks are winning the predatorial race. They are wrong. Sharks are losing, big time.

Eilperin has a refreshingly different take on how humans are contributing to the decline of shark populations around the world. Avoiding sensationalism or dry facts, she inserts herself into the daily lives of the 'shark people' who work with or on the fish. Eilperin begins her journey by digging into the psyche of the

'shark callers' of Papua New Guinea, who use rattles to lure the creatures from the deep before catching them with hand-held snares, thereby demonstrating their divine connections with the shark. She then interviews shadowy buyers of shark fins in Asia; larger-than-life shark

charter operators; the family of Peter Benchley, author of the 1974 novel *Jaws*; dedicated conservationists; and quirky shark biologists. All are portrayed in entertaining detail.

Her book draws the reader into the world of these frequently bombastic characters. One charter-boat owner who operates out of Florida boasts ▶



**Demon Fish: Travels Through the Hidden World of Sharks**  
JULIET EILPERIN  
Pantheon: 2011.  
320 pp. \$26.95



► of having killed 100,000 sharks, and specializes in hunting the biggest he can find. As he tries to satisfy the fishing needs of his clientele, you begin to see through his business-oriented eyes. Yet any appreciation soon disappears when Eilperin notes his propensity for targeting a thresher shark birthing ground, or his desire to catch and kill pregnant hammerheads because they are larger.

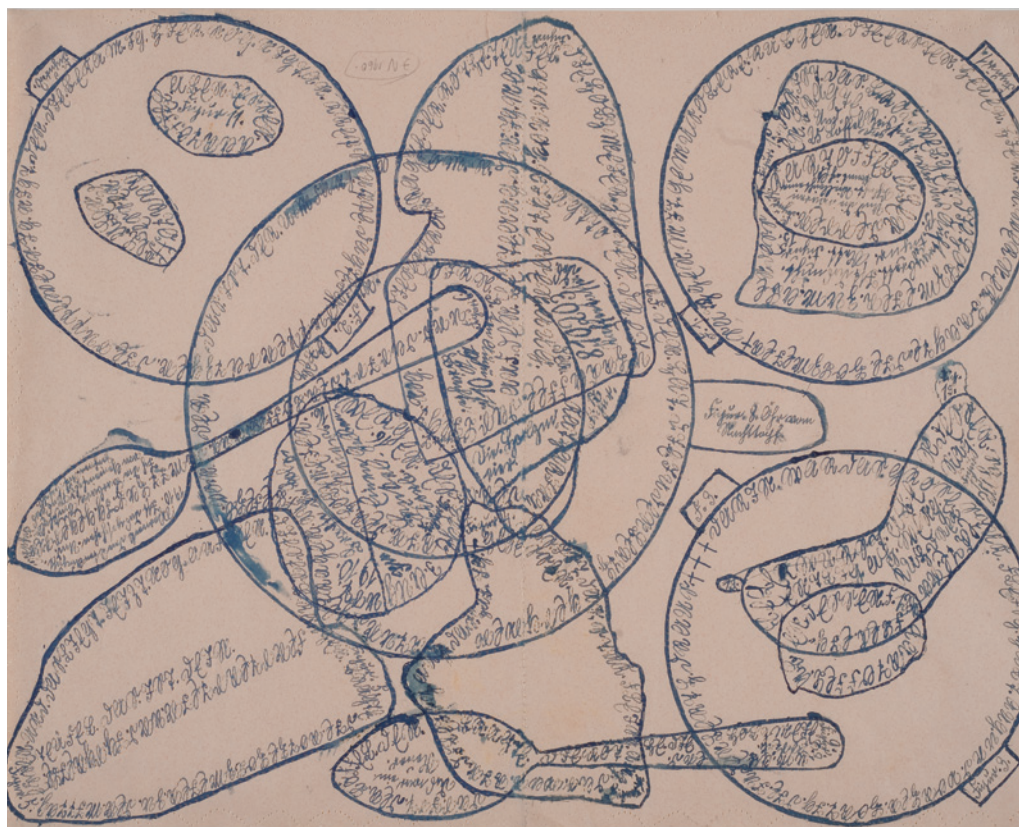
In one of the most insightful chapters — the best I've read on the subject — Eilperin describes how the burgeoning demand for shark-fin soup in Asian markets is a leading cause of the shark's global population decline. Many sharks are now caught and killed solely for the value of their fins — up to US\$2,000 per kilogram. The edible yet less-valuable carcass of the still-living shark is sometimes simply thrown back into the water. Eilperin meets elusive fin buyers and sellers, who state that if the sharks disappear from the sea altogether, "there's nothing we can do about it".

With up to \$57,000 being paid for a single basking shark fin, this is big business. Sampling the fabled Chinese wedding soup, the author pulls a tasteless fin strand out of her bowl and concludes that the expensive delicacy is a scam. The chefs she interviews agree, admitting that the fin has no culinary value. A dish that is about style, not substance, may be leading to the demise of entire species.

Although *Demon Fish* is more about people's attitudes to sharks than fish biology, Eilperin cleverly introduces science into her account. She describes the advocacy positions taken by some marine biologists attempting to stop recreational shark fishing, and relates sometimes-comical first-hand observations of biologists trying to tag sharks. Curious by its absence, however, is any wider discussion of the huge impact of accidental commercial by-catch on shark populations. The author makes the point that the most effective conservation efforts are now being made by market forces such as ecotourism, rather than through abstract environmental idealism.

From the ethics surrounding shark fishing to the delightfully creative methods of shark-conservation groups, including auctioning off naming rights for new species, *Demon Fish* captures most of the key issues affecting sharks today. ■

**Steven Campana** is a research scientist and head of the Canadian Shark Research Laboratory at the Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada.  
e-mail: [steven.campana@dfo-mpo.gc.ca](mailto:steven.campana@dfo-mpo.gc.ca)



Barbara Suckfüll's 1910 ink drawing incorporates words describing her life as a psychiatric patient.

#### PSYCHOLOGY

## Asylum art

The Prinzhorn museum shows how psychiatric patients' works have inspired artists, finds **Giovanni Frazzetto**.

**D**uring 1919–21, in a psychiatric clinic in Heidelberg, Germany, the young psychiatrist and art historian Hans Prinzhorn collected more than 5,000 works of art created by about 400 patients. Forgotten for more than 50 years after being condemned as 'degenerate art' during the Nazi regime, the collection was catalogued in the 1980s and housed in a museum in the clinic in 2001.

This year the Prinzhorn Collection Museum celebrates its tenth anniversary, and the 125th anniversary of Prinzhorn's birth, with an exhibition. A selection of original pieces from the collection are paired with works by modern artists who have drawn inspiration from them — including Paul Klee, Ernst L. Kirchner and Max Ernst.

The responses comment on the creativity of the patients, the psychiatric establishment itself and past understandings of mental pathology. Some of the reactions do not quite compete with the authenticity of the original works; others are compelling and eloquent.

**Von Kirchner bis heute (From Kirchner To Today)**  
Prinzhorn Collection  
Museum and elsewhere,  
Heidelberg, Germany.  
Until 14 August.

Drawings and an installation by contemporary artist Peter Riek revive the story of farmer Barbara Suckfüll, who, at the age of 50, started to hear voices. In 1910,

under their command, she began to draw outlines of dishes and cutlery. Writing along and in between the outlines, Suckfüll captures in words her everyday life in the Heidelberg asylum: what she thought, did or ate, her rows with the nurses and what the voices told her. Every word is followed by a full stop, resulting in a dense net of marks that dissolves into abstraction.

Also on show is a thoughtful series of drawings by Dorothee Rocke, one of the first artists to explore the collection when the museum opened in 2001. She dedicates her work to a 1912 sketch by inmate Hyacinth Freiherr von Wieser, in which

B. SUCKFÜLL OHNE TITEL (1910)/SAMMLUNG PRINZHORN



an uncertainly outlined face blends into a landscape annotated like a map. Rocke's drawings dismember the image and evolve into forms that recall elements of asylum life, such as the cells in which patients were held.

Prinzhorn saw the creations of the mentally ill, who were uninfluenced by shifting trends in the art world, as a raw depiction of an individual's condition and a valuable

way to externalize the psyche to the outer world. The exhibition is not an enquiry into the origins of mental disorders, but into their lived experience. Noticeable among the contemporary works is the absence of any reference to the brain.

At a time when psychiatry is undergoing huge change — the genetic revision of diagnoses, the search for biological markers

and the use of functional magnetic resonance imaging — this exhibition reminds us of the irreplaceable ability of personal narratives to enter the depths of the mind through doors that are not open to biological analysis. ■

**Giovanni Frazzetto** is at the BIOS Centre of the London School of Economics, UK. e-mail: g.frazzetto@lse.ac.uk

## NEUROSCIENCE

# Literary inspiration

Jorge Luis Borges' writings on memory foretell modern research, **Gabriel Kreiman** finds.

One of the most creative literary thinkers of the last century was Argentinian author Jorge Luis Borges (1899–1986). In his stories, Borges stretched the meaning and basic rules of space, time and infinity. He pondered instruments that could view the entire Universe, the effects of allowing one second to instantly become one year, infinite labyrinths and endless cyclic books. Borges was also intrigued by memory, its functions and its malfunctions.

A book in Spanish (English translation pending), *Borges y la Memoria* (*Borges and Memory*) by neuroscientist Rodrigo Quian Quiroga, eloquently links Borges' ideas with the latest research on memory. Focusing on the protagonist of Borges' 1942 essay, *Funes el Memorioso* (*Funes the Memorious*), and examining real experiments and consequences, Quian Quiroga offers a fascinating account of the connections between science and art.

The scientific study of memory often focuses on how we learn, how information is stored in circuits of neurons and how short-term memories are consolidated into long-term ones. Less attention is paid to the important ability to generalize and to forget. Borges explored both through the fictional character of Funes, who receives a brain injury after falling from a horse.

Because of his prodigious memory, Funes always knows the precise time without having to consult his watch, learns languages after merely checking a dictionary and can enumerate all the people he has ever met. Although these abilities seem admirable, there is a catch. Every leaf of every tree that he has ever seen remains in his mind. It disturbs him that a front view of a face is assigned the same identity as the profile view. He is almost incapable of understanding or creating abstract ideas.

➔ **NATURE.COM**  
For more on *Funes the Memorious*:  
[go.nature.com/5de4vm](http://go.nature.com/5de4vm)

Quian Quiroga re-examines Funes, and real people with similar amazing memory



**Borges y la Memoria**  
RODRIGO QUIAN QUIROGA  
*Sudamericana: 2011 (in Spanish). Pp65 (US\$15)*

capacities, from a modern neuroscientific perspective. We now know that the hippocampus has a crucial role in memory formation. In the 1950s, neurosurgeons observed that bilateral removal of the hippocampus for clinical purposes had profound effects on patients' abilities to establish memories. A landmark case study described patient H. M., who could easily recollect information from before his hippocampal surgery but had difficulty retaining subsequent experiences. He would meet a new person and hold a pleasant conversation, yet the next day he would have no recollection of what the person looked like or what they had spoken about.

Several important investigations ensued. Behavioural studies in rodents and

non-human primates characterized the effects of lesions in the medial temporal lobe and confirmed the central role of the hippocampus in learning. Listening in on the activity of neurons by inserting thin needles into rodents' brains revealed that hippocampal neurons can change their connection strengths in ways that correlate with learning and memory formation. Similar recordings in humans uncovered neurons in the hippocampus that showed sparse and invariant responses that are reminiscent of abstract elements of perceptions and memories.

Despite these breakthroughs, our scientific understanding of memory processes is in its infancy. We do not know how the hippocampus and cortical structures communicate to allow memory consolidation; how circuits of neurons store long-term information; whether there are capacity limits to human memory; or how to apply the properties of fault-tolerance, abstraction and speed so nicely implemented by the human brain to the design of other circuits.

Quian Quiroga's thoughtful book reminds us that elucidating the mechanisms that lead to generalization and abstraction constitutes a key step towards characterizing human intelligence. As Borges put it: "To think is to forget a difference, to generalize, to abstract". ■

**Gabriel Kreiman** is assistant professor in the Departments of Ophthalmology and Neurology at Children's Hospital Boston, Harvard Medical School, Boston, Massachusetts 02115, USA. e-mail: [gabriel.kreiman@tch.harvard.edu](mailto:gabriel.kreiman@tch.harvard.edu)



The mind's labyrinth intrigued Jorge Luis Borges.

## CORRECTION

Patricia Churchland's review (*Nature* **475**, 35; 2011) incorrectly stated that huskies run at a pace of more than 5 kilometres (3.2 miles) a minute. The pace should have been '3.2-minute miles'.

# CORRESPONDENCE

## Fishery reform: an economic problem

Rainer Froese (*Nature* 475, 7; 2011) attributes the lack of success of Europe's Common Fisheries Policy to collusion between fisheries advisers and industry. I disagree. The fundamental problem is economics.

The only way to rebuild depleted fish stocks is to catch and land fewer fish until the stocks have recovered. That means reduced earnings for fishermen, at least in the short term. Thus, any conservation measure that is likely to work is going to hurt, and is almost certain to be opposed.

The deal on offer in Europe and elsewhere is a trade-off between short-term losses and long-term gains (larger stocks leading to higher stability, catch rates and profits). Without some mechanism to make the deal more attractive to fishermen, it is always going to be a struggle to achieve sustainable levels of fishing.

Such a mechanism is conceivable: for example, through an up-front investment in a buy-back of fishing rights, to be paid for by future licence charges or resource rental taxes. These steps are necessary to prevent future pressure on conservation measures once fishing becomes more profitable again. Direct controls on fishing effort, such as limiting the number of days at sea, are also likely to be needed. They have many advantages, especially for the regrettably common situation of inadequate stock assessments, and for reducing discarding in mixed fisheries.

Economic deadlock has bedevilled fisheries management for more than half a century. It needs to be addressed directly for any progress to be made.

**John Shepherd** *National Oceanography Centre, University of Southampton, UK.*  
jgs@noc.soton.ac.uk



## Fishery reform: ban political haggling

We applaud proposals by the European Commission to reform the Common Fisheries Policy by phasing out fish discarding, broadening multi-year species-management plans, improving data collection and moving to ecosystem-based management (*Nature* 475, 7; 2011). But one vital reform has been missed: bargaining over total allowable catches should be banned and decision-makers should be compelled to follow scientific advice.

Politicians have habitually overruled scientific advice on fisheries since inception of the EC policy in the 1980s, setting total allowable catches one-third higher than recommended levels. Placing short-term political expediency and industry lobbying ahead of long-term sustainability threatens food security and the health of future generations.

Science provides the best tools for maximizing immediate benefits from fishing without squandering future opportunities. Let politicians argue for their national share of what nature can provide, rather than adopting policies that undermine the biological basis of food production. Politicians

must cede their power over fisheries if they are properly to serve the public interest.

**Bethan C. O'Leary, Callum Roberts** *University of York, UK.*  
callum.roberts@york.ac.uk

## In defence of green economy report

As contributors to the United Nations Environment Programme report criticized by Peter Victor and Tim Jackson (*Nature* 472, 295; 2011), we stand by its claim that a green economy can grow faster than a brown economy.

Victor and Jackson argue that the report's G2 green investment scenario sets out inadequate target reductions for carbon dioxide emissions by 2050. But G2 is only one path projected for the next 40 years. The assumed technologies and costs derive mainly from International Energy Agency reports. Modelling a faster transition would require stronger assumptions about future technologies.

Contrary to the authors' implication, the analysis does not assume that the funds needed to achieve this target under a green-investment scenario would exceed those under

a business-as-usual (BAU) scenario. The report explains that G2 was compared against a BAU2 scenario in which the same amount of investment (2% of global gross domestic product; GDP) is allocated to existing patterns of gross capital formation (see [go.nature.com/wdmgbb](http://go.nature.com/wdmgbb)).

Neither does the analysis overlook different living standards between countries. The T21 global model presents global totals, and it is calibrated to reflect the past 40 years, during which the ratio of the GDP per capita in high-income countries to that in low- and medium-income countries has declined from about 22:1 in 1990 (according to the World Bank) to 15:1 in 2007. Even though convergence of living standards was not set as a target, the modelled scenarios do probably represent further convergence.

We cannot claim that a green economy will always grow faster than a brown economy. But the report does provide evidence to counter the long-standing view that such an outcome is unlikely, even when limited to conventional GDP as a measure.

**Andrea Bassi** *Millennium Institute, Arlington, Virginia, USA.*  
**Derek Eaton** *United Nations Environment Programme, Geneva, Switzerland.*  
derek.eaton@unep.org

## Better provenance for biobank samples

Health-care systems spend billions of dollars annually on biomarker research for personalized medicine. Success hinges on the quality of the biobank specimens and the data used to derive them, but a lack of quality control is polluting the scientific literature with flawed information that will take a long time to sort out (see [go.nature.com/tfm8sn](http://go.nature.com/tfm8sn)).

We analysed 125 papers (see [go.nature.com/tfm8sn](http://go.nature.com/tfm8sn)).



nature.com/yagpplx) retrieved in a PubMed search of open-access articles using the key words 'biomarker discovery' for the years 2004 and 2009. More than half of these contained no information about the biospecimens used. Four papers on biomarker discoveries published in *Nature* in 2009 contained insufficient specimen data.

Leading journals are trendsetters when it comes to defining publication criteria. For example, for some 15 years they have required statements on ethical review boards and informed consent; a biomedical paper without this information would be unthinkable today. To uphold standards, all journals should insist on full details of biobanked specimens (including medication, transportation, processing and storage).

Thousands of potential biomarkers are reported every year. Biobankers should collect complete information on specimens and pass it on to researchers to include in their publications.

**Daniel Simeon-Dubach, Aurel Perren** *Biobank-Suisse, Bern, Switzerland.*  
daniel.simeon-dubach@biobank-suisse.ch

## Some more biofuel lessons from Brazil

Marcia Moraes highlights lessons from Brazil's experience as a biofuel producer (*Nature* **474**, S25; 2011), but some important ones are missing if production is to meet social and environmental standards.

First, treat the workforce consistently, and well. In some regions, sugar-cane workers are well paid, but in others — notably in west-central and northeastern Brazil — they do not have proper contracts and conditions are appalling.

Second, include small-scale farmers, who have been excluded from Brazil's ethanol sector since it started 35 years ago, on the grounds of economic inefficiency. Follow the example of its national programme for biodiesel, which promotes partnerships between

the oilseed industry and small farmers and peasants.

Third, follow environmental laws. Sugar-cane farms preserve only 11% of their land as forest (Brazilian Institute of Geography and Statistics *Censo Agropecuário 2006* IBGE; 2009), although the Forest Code legislation stipulates that at least 20% should be forested. Even if the Brazilian government revises the law to reduce that percentage, sugar-cane farms will not have sufficient natural reserves to meet the legislation.

Despite much opposition, sugar-cane production has already been prohibited in the Amazon and Pantanal regions, even while deforestation proceeds unabated in the biodiversity hot spots of the Cerrado biome and the mesophytic Atlantic rainforest.

**Rafael Feltran-Barbieri, Ricardo Abramovay, Jean Paul Metzger** *University of Sao Paulo, Brazil.*  
rafaelfb@usp.br

## Monitoring ash clouds for aviation

To minimize airport closures caused by volcanic ash, aviation authorities should make more use of methods that are now available for measuring airborne particles.

Optical particle-counter (OPC) instruments were used to measure dust pollution during Iceland's Grímsvötn volcano eruption in May this year, after a trial last year in Eyjafjallajökull's ash plume. The data were used by Isavia, the country's air-navigation services provider, to help in decision-making.

OPC instruments measure airborne concentrations of three different size classes of particulate matter components and total suspended particles, every 6 seconds. The instruments were mounted on a light, slow-flying, piston-engine aircraft after calibration in wind-tunnel experiments with volcanic ash.

As Grímsvötn erupted, OPC results were quickly dispatched to the local meteorological office and to Isavia. This helped to keep Keflavik International Airport

open for longer than expected, resulting in large savings for the airline companies.

Other European aviation authorities are considering similar operations. Also, universities in Iceland and Japan are considering using OPC measurements to improve modelling of ash-cloud events.

**Jónas Eliasson** *University of Iceland, Reykjavik, Iceland.*  
**Asgeir Pálsson** *Isavia, Reykjavik Airport, Iceland.*  
**Konradin Weber** *Düsseldorf University of Applied Sciences, Germany.*  
konradin.weber@fh-duesseldorf.de

## Humility needed in decision-making

Daniel Gilbert maintains that people generally make bad decisions on risk issues, and suggests that communication strategies and education programmes would help (*Nature* **474**, 275–277; 2011). This version of the deficit model pervades policy-making and branches of the social sciences.

In this model, conflicts between expert and public perceptions of risk are put down to the difficulties that laypeople have in reasoning in the face of uncertainties rather than to deficits in knowledge per se. There are three problems with this stance.

First, it relies on a selective reading of the literature. Evolutionary psychologists have a more positive view of people's capacity for statistical reasoning (see, for example, L. Cosmides and J. Tooby *Cognition* **58**, 1–73; 1996), arguing that many putative reasoning 'errors' may be nothing of the sort.

Second, it rests on some bold extrapolations. For example, it is not clear how the biases Gilbert identifies in the classic 'trolley' experiment play out in the real world. Many such reasoning 'errors' are mutually contradictory — for example, people have been accused of both excessive reliance on and neglect of generic 'base-rate' information to judge the probability of an event. This casts doubt on the idea that they reflect

universal or hard-wired failings in cognition.

The third problem is the presentation of rational choice theory as the only way of deciding how to handle risk issues. Alternative decision logics, from the precautionary principle to deontology, are reduced to mere reasoning fallacies. Yet to be concerned with fundamental rights, moral obligations (deontology) and worst-case scenarios (precaution) is not pathological. To treat it as such, as Gilbert and others do (see, for example, C. R. Sunstein *Behav. Brain Sci.* **28**, 531–542; 2005), seems myopic.

Given that many modern risk crises stem from science's inability to foresee the dark side of technological progress, a little humility from the rationality project wouldn't go amiss.  
**Brian H. MacGillivray, Nick F. Pidgeon** *Cardiff University, Cardiff, UK.*  
macgillivraybh@cardiff.ac.uk

## Tweet to collaborate with poorer nations

Romain Murenzi proposes ways to improve opportunities for young scientists in the developing world (*Nature* **474**, 543; 2011). In addition to these, social media could help to consolidate relationships with colleagues in technologically more advanced countries, and in greater numbers than travelling fellowships allow.

Sites such as Facebook and Twitter were instrumental in this year's topplings of North African dictatorships. They have also been used by Western agencies to support postgraduate education of medical interns in the Middle East and sub-Saharan Africa (A. E. T. Finlayson *et al. J. Telemed. Telecare.* **16**, 181–184; 2010).

A network of trusted scientists using these tools could mentor ambitious investigators in regions where free speech and unrestrained critical inquiry are novelties.

**Faisal R. Ali** *University of Manchester, UK.*  
f.r.ali.01@cantab.net  
**Alexander E. T. Finlayson** *University of Oxford, UK.*

## An icon knocked from its perch

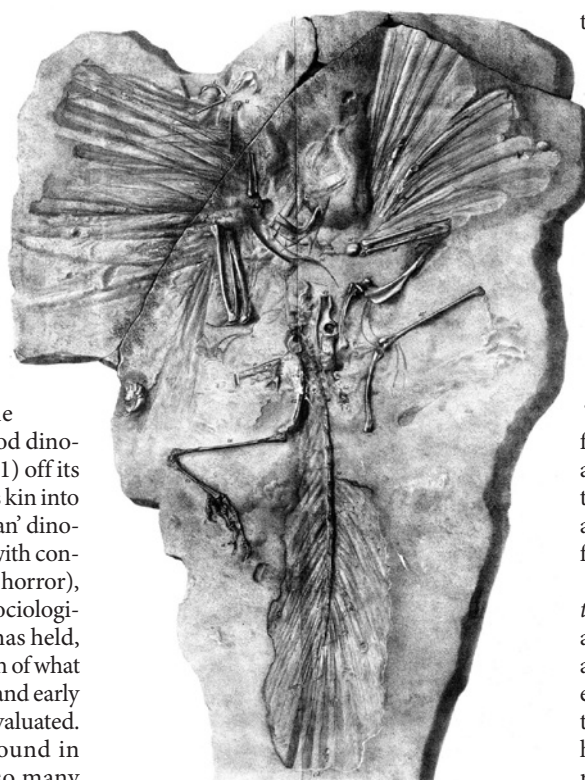
As sesquicentennial celebrations commemorate the discovery of *Archaeopteryx* as a historical symbol of evolution and the oldest fossil bird, new work shakes the dinosaur family tree — and our view of this icon. [SEE ARTICLE P.465](#)

LAWRENCE M. WITMER

It's been a good run for *Archaeopteryx*. For the past 150 years, the famous feathered fossil species from Bavaria in Germany has been a symbol of evolution, a textbook example of a transitional fossil and, above all, the oldest and most primitive bird. On page 465 of this issue, however, Xu and colleagues<sup>1</sup> present a newly discovered *Archaeopteryx*-like species named *Xiaotingia zhengi* that rearranges the branches on the phylogenetic tree of bird-like theropod dinosaurs, knocking *Archaeopteryx* (Fig. 1) off its celebrated perch and moving it and its kin into the great unwashed ranks of 'non-avian' dinosaurs. This finding is likely to be met with considerable controversy (if not outright horror), in part because of the historical and sociological significance that *Archaeopteryx* has held, but also because it may mean that much of what we thought we knew about the origin and early evolution of birds will need to be re-evaluated.

The fossils of *Xiaotingia* were found in Liaoning Province, China, where so many other spectacular specimens of feathered dinosaurs and early birds have been discovered<sup>2</sup>. The precise provenance of the specimen is a little uncertain, because it was purchased from a dealer, but all indications are that it comes from the Tiaojishan Formation, which dates to the Late Jurassic, some 155 million years ago. The chicken-sized skeleton, splayed out on a slab of shale and surrounded by a halo of feather impressions, isn't as striking as the specimens of *Archaeopteryx*, but its constellation of subtle bony bumps and grooves makes *Xiaotingia* a game-changer.

The main players here are an assemblage of bird-like dinosaurs — oviraptorosaurs and deinonychosaurs (including troodontids and dromaeosaurids) — and dinosaur-like birds that belong to the avialans. Deinonychosaurs and avialans together comprise a group known as Paraves, with oviraptorosaurs being a bit more distantly related (Fig. 2). As more fossils of basal members of each of these groups have been collected, distinctions between the groups have predictably blurred, and some



**Figure 1 | *Archaeopteryx*.** For 150 years, the Jurassic fossils of *Archaeopteryx* have been a textbook example of an evolutionary transition. From Plate 1 in ref. 11.

species have bounced around from group to group. For example, *Anchiornis*, which recently made headlines for its colourful plumage<sup>3</sup>, was originally regarded as a basal avialan<sup>4</sup>, then a basal troodontid<sup>5</sup>, and is now considered an archaeopterygid<sup>1</sup>.

Enter *Xiaotingia*. When Xu *et al.*<sup>1</sup> ran a phylogenetic analysis combining the attributes of *Xiaotingia* with those of *Archaeopteryx*, other basal avialans, deinonychosaurs, and oviraptorosaurs, not only did *Xiaotingia* and *Anchiornis* cluster with *Archaeopteryx*, but these archaeopterygids now were yanked out of Avialae and placed in Deinonychosauria (Fig. 2). In other words, *Archaeopteryx* was no longer a bird. Surprised by this outcome, the authors re-ran the analysis with identical parameters, but this time omitting *Xiaotingia*. The result was that *Archaeopteryx* was restored

to Avialae as the most basal bird. This experiment affirmed how crucial *Xiaotingia* is to understanding the evolution of advanced theropods.

It may seem heretical to say that *Archaeopteryx* isn't a bird, but this idea has surfaced<sup>6</sup> occasionally since as far back as the 1940s. G. S. Paul has been the most vociferous advocate, even going so far as to make dromaeosaurs a subfamily within the Archaeopterygidae<sup>7</sup>, thus moving *Archaeopteryx* well outside the birds. Moreover, there has been growing unease about the avian status of *Archaeopteryx* as, one by one, its 'avian' attributes (feathers, wishbone, three-fingered hand) started showing up in non-avian dinosaurs. Perhaps the time has come to finally accept that *Archaeopteryx* was just another small, feathered, bird-like theropod fluttering around in the Jurassic.

But why is this such a big deal? *Archaeopteryx* has always been something of a celebrity and has monumental historical, sociological and even political importance. It was discovered, with perfect timing, in mid-1861, less than two years after Darwin's *Origin of Species* hit the bookshelves<sup>8</sup>. With its blend of avian and reptilian characteristics (not to mention the charismatic beauty of the fossils themselves), *Archaeopteryx* was seemingly the ideal evolutionary intermediate, instantly entering the debates over evolution in Victorian England and elsewhere, and gaining prominence in textbooks.

Given this iconic role, *Archaeopteryx* has also been in the cross-hairs of creationists, and remains a lightning rod for political debates and legal proceedings about teaching evolution in schools. Of course, Xu and co-workers' finding only deepens the impact of *Archaeopteryx* by highlighting the rich evolutionary nexus of which it is a part, but how the ever-clever creationist community will 'spin' it remains to be seen.

Politics aside, the historical importance of *Archaeopteryx* stands, even if we need to add the footnote that current evidence no longer regards it as the oldest bird. The impact of losing *Archaeopteryx* from the avian clan is, nevertheless, likely to rock the palaeontological community for years to come simply because, for the past century and a half,

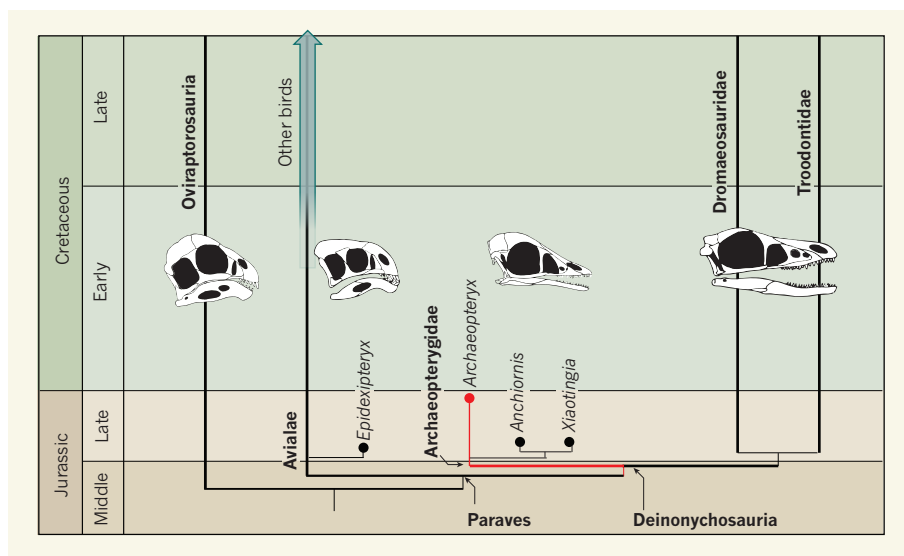


these familiar fossils have guided almost all scientific thought about the beginnings of birds. The late John Ostrom, the most influential modern worker on avian origins, began his seminal 1976 article<sup>9</sup>, itself entitled '*Archaeopteryx* and the origin of birds', with the statement, "The question of the origin of birds can be equated with the origin of *Archaeopteryx*, the oldest known bird." Indeed, virtually all our notions about early avian evolution have been viewed through the lens of *Archaeopteryx*. Hundreds of publications (including several of my own) draw on the structure of *Archaeopteryx* to formulate and evaluate hypotheses about birds. Some published phylogenetic analyses have even used *Archaeopteryx* as the sole representative of birds. To what species do we now turn to ground our understanding of early birds?

According to Xu and colleagues' analysis<sup>1</sup>, the most basal fossil birds are forms such as *Epidexipteryx*, *Jeholornis* and *Sapeornis*, all of which were named in the past decade and so comprise new territory even for specialists. Clearly, without the safety net of good old *Archaeopteryx* at the base of the birds, we've got some fresh work to do.

Once we stop whining, of course, we can see that this exciting finding actually resolves some incongruities. For example, recent work<sup>10</sup> suggests that herbivory may have been common among advanced bird-like dinosaurs, with carnivory potentially being secondarily evolved in the deinonychosaurs. Moving archaeopterygids and their carnivorous skulls out of birds and into the carnivorous deinonychosaur group makes the herbivorous oviraptorosaur-like skulls of basal birds more consistent with this new hypothesis of widespread herbivory (Fig. 2).

In truth, this chapter of the scientific story is just beginning. Just as *Xiaotingia* moved *Archaeopteryx* out of the birds, the next find could move it back in — or to somewhere else within this fuzzy tangled knot that makes up



**Figure 2 | A cluster of feathered dinosaurs.** *Archaeopteryx* has historically been regarded as the most basal bird (avialan), but the discovery of the closely related *Xiaotingia* led Xu *et al.*<sup>1</sup> to pull these archaeopterygids out of avialans (birds) and into deinonychosaurs along with dromaeosaurids and troodontids. This new grouping better accounts for the evolution of feeding strategies among bird-like dinosaurs. Previous research<sup>10</sup> suggested that herbivory was common among this group, as reflected in the tall, boxy skulls of oviraptorosaurs and basal avialans such as *Epidexipteryx*. The triangular, sharp-toothed skull of *Archaeopteryx* was incongruous among basal avialans, but fits better among the carnivorous dromaeosaurids and troodontids.

the origins of birds and bird-like dinosaurs. That said, during this sesquicentennial anniversary of *Archaeopteryx*, which is being honoured with exhibits and commemorative coins, the bitter irony may be that it may not have been the bird we've always thought it was. But *Archaeopteryx* will remain an icon of evolution, perhaps even more so now, providing compelling evidence that, as we should expect, evolutionary origins are rather messy affairs. ■

**Lawrence M. Witmer** is in the Department of Biomedical Sciences, Heritage College of Osteopathic Medicine, Ohio University, Athens, Ohio 45701, USA.

e-mail: witmerL@ohio.edu

1. Xu, X., You, H., Du, K. & Han, F. *Nature* **475**, 465–470 (2011).
2. Norell, M. A. & Xu, X. *Annu. Rev. Earth Planet. Sci.* **33**, 277–299 (2005).
3. Li, Q. *et al. Science* **327**, 1369–1372 (2010).
4. Xu, X. *et al. Chinese Sci. Bull.* **54**, 430–435 (2009).
5. Hu, D., Hou, L., Zhang, L. & Xu, X. *Nature* **461**, 640–643 (2009).
6. Witmer, L. M. in *Mesozoic Birds* (eds Chiappe, L. M. & Witmer, L. M.) 3–30 (Univ. California Press, 2002).
7. Paul, G. S. *Predatory Dinosaurs of the World* (Simon & Schuster, 1988).
8. Wellnhofer, P. *Archaeopteryx: the Icon of Evolution* (Verlag Dr. Friedrich Pfeil, 2009).
9. Ostrom, J. H. *Biol. J. Linn. Soc.* **8**, 91–182 (1976).
10. Zanno, L. E. & Makovicky, P. J. *Proc. Natl Acad. Sci. USA* **108**, 232–237 (2011).
11. Owen, R. *Phil. Trans. R. Soc. Lond.* **53**, 33–47 (1863).

## PRECISION MEASUREMENT

# Exciting antiprotons

**Shining light on an antiproton masquerading as an electron in a helium atom is a rich source of physics. An approach that brings this technique to unprecedented precision will allow new tests of matter–antimatter symmetry. SEE LETTER P.484**

MIKE CHARLTON

In a 1991 News & Views article<sup>1</sup> on the ingenious means by which positrons, the antiparticle counterpart of electrons, were used in materials science, John Maddox, the then editor of *Nature*, asked: "Can any antiparticle make a probe?" What prompted the

question was the discovery<sup>2</sup> that a few per cent of the antiprotons injected and stopped in liquid helium were trapped in long-lived (metastable) states, with lifetimes of several microseconds. Twenty years down the line, Maddox's poser can be answered with an unequivocal 'yes'. The three-body atomic system that is formed when an antiproton replaces one

of the electrons in a helium atom has proved a superb probe for fundamental physics.

In the latest advance, Hori and co-workers<sup>3</sup> (page 484) describe a trick that has enabled transitions between selected energy levels of the antiprotonic helium to be measured with unprecedented precision. The bottom line is that these measurements, when combined with theory<sup>4</sup>, have allowed, among other things, a more accurate value of the antiproton-to-electron mass ratio to be determined in order to compare it with the proton equivalent.

The source of the antiprotons is the Antiproton Decelerator located at CERN, the European particle-physics laboratory near Geneva in Switzerland. This unique machine provides pulses of antiprotons every hundred seconds or so. After deceleration, the antiparticles are injected into a cell containing helium

gas, where the metastable atoms are formed. When the antiproton removes an electron, it occupies an orbital with roughly the same binding energy as that of the electron, but one that has a much higher principal quantum number. The transition energies to nearby states are in the optical or near-optical regions, and so are amenable to probing using lasers. A short time after the arrival of the antiprotons, laser pulses are fired at the atoms to drive the antiprotons into a state that does not have the metastable property, whereupon they promptly annihilate on contact with the helium nucleus. The particles, known as pions, that are ejected as a result of this annihilation are tremendously energetic and can easily be detected to find the resonant condition at which the laser frequency matches that of the transition.

Such investigations have been feasible for some time now<sup>5</sup>, so what is new about this one? One of the main aims of atomic spectroscopy is to measure spectral-line frequencies as accurately as possible, and one of the effects that need to be overcome is that caused by the motion of the atoms. The change in the frequency of the sound emitted from a moving vehicle (say) as it passes us is a commonplace experience, and is known as the Doppler shift. A similar effect also applies to the absorption of light by atoms. However, because the atoms move in random directions with respect to the laser beam, the shifts are all different, but nonetheless they have the effect of smearing out the response of the atom. This is called Doppler broadening.

One way to reduce, or even eliminate, Doppler broadening is to excite the transition using two photons rather than one — an effect used to achieve staggering levels of precision in the spectroscopy of atomic hydrogen<sup>6</sup>. Two counter-propagating laser beams of the same frequency are shone on an atomic ensemble; if an atom simultaneously absorbs one photon from each beam, then the Doppler broadening is negated.

This approach works well for exciting electrons, but antiprotons are much more sluggish, and to excite them in this way would require colossal laser power. However, the effects of Doppler broadening can be partially overcome by using photons of different frequencies, and arranging that one of the photons drives the antiproton close to, but not actually to, a real atomic state. This 'sleeping' state acts as a kind of helping hand to the antiproton, making the two-photon transition feasible. It turns out that the reduction in the broadening is given by a factor of the sum of the photon frequencies, divided by their difference. In the case of the experiment of Hori and colleagues<sup>3</sup>, this factor was around 10. The dramatic narrowing of the antiprotonic helium lines is clearly visible by comparing the two-photon signal with its single-photon counterpart (see Fig. 2a and b on page 485).

The value for the antiproton-to-electron mass ratio obtained by the authors<sup>3</sup> was

1836.1526736(23), where the numbers in parentheses are the uncertainties on the last two digits. This compares well with the corresponding value for the proton. Taking their analysis further, Hori *et al.* were able to improve the comparisons of the charge and mass of the antiproton with those of the proton by 10,000-fold, reaching a precision just below one part in a billion.

This is impressive progress. But why are studies like these important? The main reason is that our knowledge of the symmetries that underpin the laws of nature is incomplete. At present we are very far from understanding what happened to all the antimatter that was created in equal proportion to matter in the Big Bang. The Universe, it seems, is made up from the remnant matter — but we don't know how this came about. Seeking the answer to this conundrum is an active area of science, encompassing the Large Hadron Collider, cryogenic antiprotons and cosmic-ray physics.

Probing interactions at low energies to look

for small deviations from expected results (see, for example, refs 7 and 8) is a painstaking and challenging discipline. But if any effect, such as a deviation in the ratios studied here, were found, no matter how small, the impact on physics would be dramatic. Recent progress has been encouraging, but much further effort is required. The search for nature's defects goes on. ■

**Mike Charlton** is in the Department of Physics, College of Science, Swansea University, Swansea SA2 8PP, UK.  
e-mail: m.charlton@swansea.ac.uk

1. Maddox, J. *Nature* **353**, 207 (1991).
2. Iwasaki, M. *et al. Phys. Rev. Lett.* **67**, 1246–1249 (1991).
3. Hori, M. *et al. Nature* **476**, 484–488 (2011).
4. Korobov, V. I. *Phys. Rev. A* **77**, 042506 (2008).
5. Hayano, R. S. *et al. Rep. Prog. Phys.* **70**, 1995–2065 (2007).
6. Niering, M. *et al. Phys. Rev. Lett.* **84**, 5496–5499 (2000).
7. Pohl, R. *et al. Nature* **466**, 213–216 (2010).
8. Hudson, J. J. *et al. Nature* **473**, 493–496 (2011).

#### ASTRONOMY

## A census of the Carina complex

**Dust and glowing hydrogen obscure the Carina complex at visible wavelengths. An X-ray study, combined with infrared surveys, provides knowledge of newly formed stellar associations and past supernova explosions in this system.**

THEODORE R. GULL

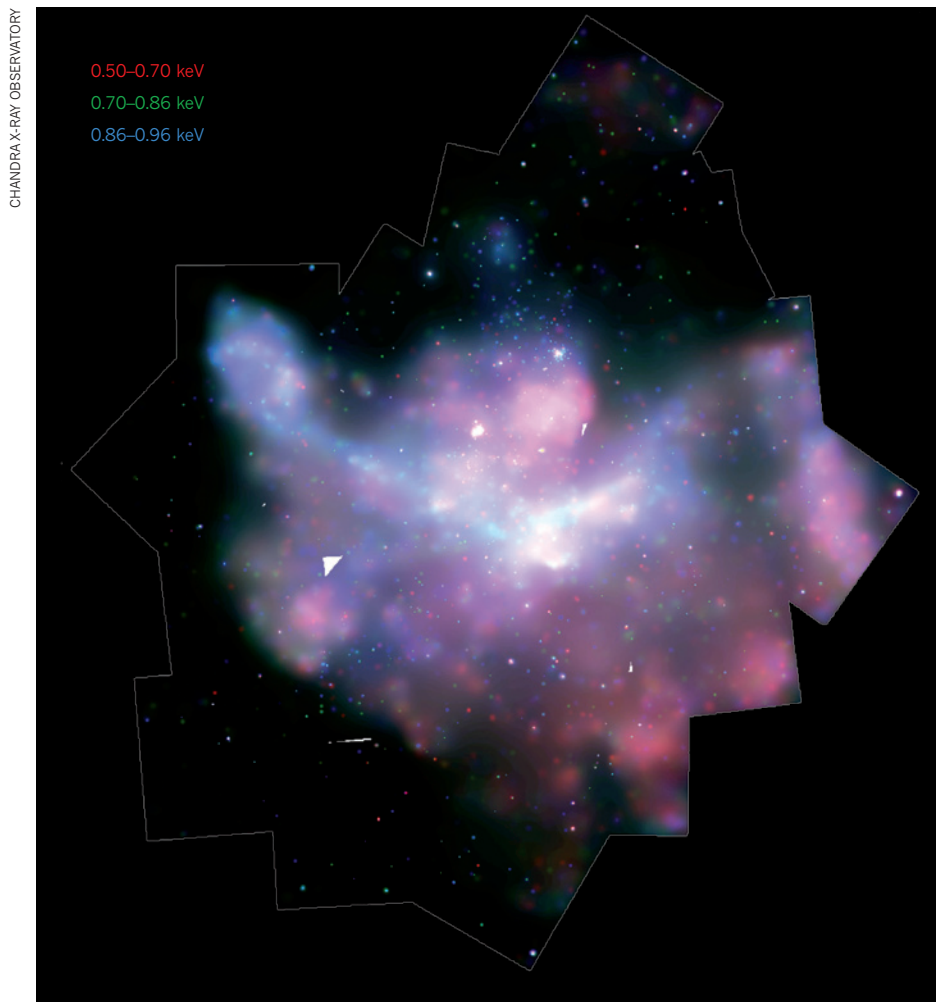
**A**n association of young stars and its nascent environment is a chaotic region driven by massive stellar winds and ejecta from explosions known as supernovae. Whereas the region's diffuse core may glow with light from the reformation of hydrogen that had been ionized by the hottest stellar members, its dusty cloud hides the cooler stellar membership. At visible wavelengths, a full inventory of such a recently formed (within the past 10 million years) stellar group is nearly impossible. Multi-wavelength studies, from radio to infrared to ultraviolet and even X-rays, are essential to gaining a complete picture of how a stellar association forms. A large team of astronomers has now published a string of sixteen papers, including four<sup>1–4</sup> discussed here, describing an X-ray study that sheds light on one of the most enigmatic of such chaotic regions — the Carina complex. The papers form an entire issue of the *Astrophysical Journal Supplement Series*.

Picture an area in the sky, located in the Milky Way, that is three times the diameter of the Moon. Viewing the Galaxy's

Sagittarius–Carina spiral arm from Earth, astronomers find within the Carina region an extraordinary complex of massive stars, recently formed stellar associations and large regions of ionized hydrogen bounded by dusty clouds that have ongoing stellar birthplaces buried within. The obscuring dust limits astronomers from obtaining a full understanding of everything that is going on in this region. But as the Galaxy is relatively transparent to X-rays, the X-ray spectral region provides a means of peering through much of this dust and gaining a more complete census.

In their study, known as the Chandra Carina Complex Project, the team of astronomers obtained a deep X-ray survey of the Carina complex (Fig. 1) using the Chandra X-ray Observatory, one of the four great observatories launched by NASA in the past two decades. The team's focus was on point-like X-ray sources, as well as diffuse X-ray emission, from recent and ongoing star-formation regions within the Carina complex. The Chandra Observatory provided unsurpassed X-ray sensitivity and spatial resolution, allowing separation of extended diffuse emission from point-like





**Figure 1 | Carina complex in focus.** This image, spanning about three lunar diameters on the sky, shows X-ray emission from the Carina complex in three energy bands. The energy ranges of the bands, given in kiloelectronvolts (keV), have been assigned, in ascending order of energy, the colours red, green and blue. White triangles and other white blemishes are unobserved sky areas. Image modified from ref. 5.

sources. Over 14,000 point sources were detected within a 1.5-square-degree region of sky observed in the allocated 1.2 megaseconds of observatory time. In comparison, the entire catalogue of X-ray sources found by the Einstein Observatory in the 1980s, obtained with much lower spatial resolution and X-ray sensitivity, listed only 5,000 sources scattered over an extensive portion of the celestial sphere. Within the Carina complex, only a few dozen sources were identified.

The extended team compared the X-ray data to visible and infrared data obtained from ground and space observatories. Positional correlations between the point sources in the several surveys, along with their combined spectral energy distributions, led to the identification of the X-ray sources. Gagné *et al.*<sup>1</sup> demonstrated that, of the 200 known O- and B-type stars (the two brightest and most massive class of stars), 68 of the 70 O-stars and about half of the earlier (more massive, hotter) B-type stars were detected at X-ray wavelengths.

There are at least four mechanisms of X-ray emission from hot massive stars in general.

First, single O-stars have such massive winds that they develop instabilities, leading to density clumps within the winds. High-speed gas moves past slow-moving gas clumps, resulting in shocks and X-ray emission. Second, the massive winds of binary O-star systems in close, elliptical orbits collide, leading to strong X-ray emission, which is modulated by the stars' orbital period (the famous Eta Carinae binary is a classic example of this phenomenon). Third, a few O-stars have high, large-scale magnetic fields (with a strength of about 1 kilogauss) resulting in gas flows from their polar regions that collide at their magnetic equators, thus leading to strong X-ray emissions. Finally, some B-stars seem to have lower-mass companions at the 'pre-main-sequence' stage of stellar evolution; sometimes, their X-ray emissions are the only indications of the unseen companions. More results from the Chandra Carina Complex Project will lead to greatly increased knowledge of these mechanisms and the possible discovery of new causes of X-ray emission in massive stars.

Povich *et al.*<sup>2</sup> compared the X-ray point sources to the Two Micron All-Sky Survey and

the Spitzer Space Telescope Vela-Carina survey and identified 94 candidate O- or B-stars that now await confirmation by subsequent studies, thus potentially increasing the number of known OB stars in the Carina region by 50%. Preibisch *et al.*<sup>3</sup> made a partial study of the Carina complex at sub-arcsecond angular resolution in the near infrared using the European Southern Observatory Very Large Telescope. Of the more than 7,000 Chandra X-ray sources within the surveyed field, nearly 90% have an identified near-infrared stellar counterpart. (Preibisch and colleagues' study is sensitive to near-infrared stellar counterparts down to 0.5 to 1.0 solar masses.) Many more results will be published in the near future, most notably on the stellar initial mass function (how many stars formed for each mass interval) of the Carina stellar associations.

Perhaps the most important result of the Chandra Carina Complex Project is the determination of the diffuse X-ray contribution by Townsley and colleagues<sup>4</sup>. The team carefully subtracted out all point-like (stellar) sources and found that the majority of the X-ray emission is truly diffuse. Much of the X-ray emission can be associated with the edges of cold, neutral clouds where fast-moving gas in the winds and supernova ejecta collide. The massive wind of a single O-star is known to create an interstellar bubble that grows as the star evolves. Eventually the O-star explodes as a supernova that increases the size of the bubble. With time, multiple bubbles from many O- and B-star winds and supernovae overlap and combine into a superbubble. Indeed, the Carina nebula — the extended region of emission — is a superbubble, being the combination of multiple stellar winds and supernova ejecta. Across the surveyed region, Townsley and colleagues' analysis of the diffuse X-ray emission allowed them to identify portions of the X-ray nebula as caused by stellar winds (low-energy X-rays) and supernova ejecta (high-energy X-rays).

This impressive X-ray survey and its initial correlations with visible and infrared surveys open a treasure trove that will lead to much new understanding on how star associations form and evolve with time. We anticipate many, even more exciting, results to appear in the journals soon. ■

**Theodore R. Gull** is in the Astrophysics Science Division, Goddard Space Flight Center, NASA, Greenbelt, Maryland 20771, USA.  
e-mail: ted.gull@nasa.gov

1. Gagné, M. *et al. Astrophys. J. Suppl. Ser.* **194**, 5 (2011).
2. Povich, M. S. *et al. Astrophys. J. Suppl. Ser.* **194**, 6 (2011).
3. Preibisch, T. *et al. Astrophys. J. Suppl. Ser.* **194**, 10 (2011).
4. Townsley, L. K. *et al. Astrophys. J. Suppl. Ser.* **194**, 15 (2011).
5. Townsley, L. K. *et al. Astrophys. J. Suppl. Ser.* **194**, 1 (2011).

# Magnesium in a signalling role

Magnesium binds to enzymes and nucleic acids and is essential for their activity. It emerges that this ion can also function as a signalling molecule with a crucial role in the immune system. [SEE ARTICLE P.471](#)

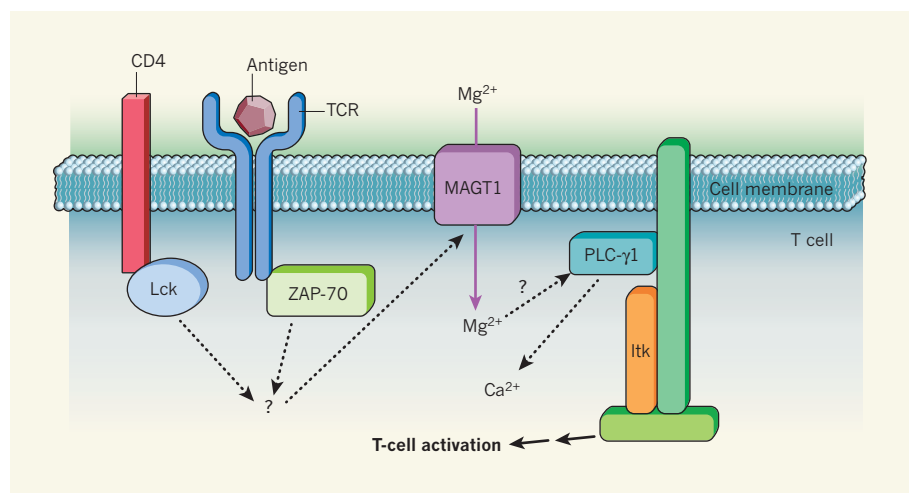
NING WU & ANDRÉ VEILLETTE

Humans fend off offending agents such as pathogens and tumour cells using their immune system. When immune function is diminished as a result of immunodeficiencies, individuals become more susceptible to infections and, in some cases, to cancer. Immunodeficiencies can result from defects in any component of the immune system, including T cells and B cells. These disorders can be hereditary (primary immunodeficiencies) or acquired — for example, through infection with viruses such as HIV<sup>1</sup>. Although primary immunodeficiencies are rare, their study has frequently led to identification of new, and at times unsuspected, regulators of the immune system. On page 471 of this issue, Li and colleagues<sup>2</sup> describe one such regulator, the magnesium-ion transporter MAGT1, mutations in which result in a hereditary human immunodeficiency.

T cells recognize and react to foreign antigens through T-cell receptors (TCRs) on their surface<sup>3</sup>. Antigen binding to these receptors initiates a signalling cascade, which

activates the T cell, leading to a protective immune response. The cascade involves kinase enzymes, such as Lck, ZAP-70 and Itk, which, among other substrates, phosphorylate phospholipase C- $\gamma$ 1 (PLC- $\gamma$ 1). This enzyme in turn regulates intracellular levels of calcium — a crucial inducer of T-cell activation. Mutations in several proteins involved in TCR signalling, including ZAP-70 and Itk, have been discovered in patients with primary immunodeficiencies<sup>1</sup>.

Li *et al.*<sup>2</sup> studied three male patients from two different families who had a new type of primary immunodeficiency. In one well-characterized family, two brothers showed chronic infection with the Epstein–Barr virus (the cause of infectious mononucleosis) from a young age. Both also had abnormally low numbers of CD4<sup>+</sup> T cells (a subset of T cells involved in antiviral immune responses and in the control of antibody production by B cells) and showed defects in TCR signalling. The third patient had a related disorder, although the authors could not examine him as thoroughly as the others because of his untimely death.



**Figure 1 | MAGT1 in T-cell activation.** Antigen binding to a T-cell receptor (TCR) results in a series of molecular changes, initiated by the kinase enzymes Lck (bound to the CD4 receptor), ZAP-70 and Itk. Among the proteins phosphorylated by these kinases are phospholipase C (PLC)- $\gamma$ 1, which regulates intracellular levels of calcium ions ( $\text{Ca}^{2+}$ ). TCR stimulation also results in activation of the ion transporter MAGT1, leading to an increase in intracellular levels of free magnesium ions ( $\text{Mg}^{2+}$ ). The intermediate steps linking TCR stimulation to MAGT1 activation are not known. The MAGT1-mediated increase in free  $\text{Mg}^{2+}$  augments T-cell activation at least in part by promoting activation of PLC- $\gamma$ 1.

Genetic analyses revealed that all three patients had mutations that abolished the function of the *MAGT1* gene on their X chromosome. The gene's location explains why only males, which have a single copy of the X chromosome, were afflicted by the disorder. Females carrying this mutation on one of their X chromosomes usually do not show any symptoms — as was the case for the mother of the two affected brothers — but can transmit the mutations to their offspring.

The protein product of *MAGT1* is the MAGT1 transporter in the cell membrane that allows magnesium ions ( $\text{Mg}^{2+}$ ) to flow into the cell<sup>4</sup>. Most of the  $\text{Mg}^{2+}$  in a cell is bound to DNA, to RNA, to the cellular energy carrier ATP or to enzymes, and acts as an essential cofactor for these molecules. Nonetheless, a small fraction (less than 5%) occurs as free  $\text{Mg}^{2+}$ , and the role of this pool is poorly understood.

Li *et al.* show that in *MAGT1*-deficient patients, T cells cannot effectively increase their intracellular free  $\text{Mg}^{2+}$  levels in response to TCR stimulation. These cells also showed a decrease in PLC- $\gamma$ 1 activity and in the amount of intracellular calcium (Fig. 1). Despite these defects in T-cell activation (and abnormalities in certain responses in epithelial cells), the B cells in these patients remained unaffected.

Previous work<sup>5</sup> has implicated free  $\text{Mg}^{2+}$  in T-cell activation. Li and colleagues' work now reveals the physiological relevance and the likely mechanism of this effect. Many questions remain, however. For instance, are defects in T-cell activation in *MAGT1*-deficient patients solely due to a problem with  $\text{Mg}^{2+}$  regulation? It could be that *MAGT1* has other — as yet unidentified — functions, defects in which contribute to immunodeficiency.

Other questions include exactly how free magnesium influences PLC- $\gamma$ 1 and calcium levels in T cells. And are its effects direct, or are they indirect, occurring through another protein such as the kinase Itk? Li *et al.* also find that B cells do not seem to increase their intracellular levels of free  $\text{Mg}^{2+}$  in response to activation, even though these cells express *MAGT1*. Can this difference be explained by variations between the signalling machineries of B cells and T cells? Or are there, perhaps, other  $\text{Mg}^{2+}$  regulators that substitute for *MAGT1* activity in B cells?

It will be interesting to determine which cell types — other than T cells and epithelial cells — are affected by the absence of *MAGT1*. A way to explore this might be to investigate whether patients deficient in this transporter show clinical manifestations other than T-cell immunodeficiency. Finally, could drugs be developed to suppress the activity of *MAGT1*, or of the targets of free  $\text{Mg}^{2+}$  within the cell, to control excessive T-cell-mediated immune responses such as those seen in human autoimmune diseases including diabetes, lupus



and rheumatoid arthritis, or in transplant rejection? These questions will surely intrigue scientists and clinicians alike. ■

**Ning Wu and André Veillette** are in the Laboratory of Molecular Oncology, Clinical Research Institute of Montreal, Montreal, Quebec, H2W 1R7 Canada. André Veillette is also in the Department of Medicine, University of Montreal, and in the Department of

Medicine, McGill University, Montreal.  
e-mail: [veillea@ircm.qc.ca](mailto:veillea@ircm.qc.ca)

1. Casanova, J. L. & Abel, L. *Science* **317**, 617–619 (2007).
2. Li, F. *et al.* *Nature* **475**, 471–476 (2011).
3. Smith-Garvin, J. E., Koretzky, G. A. & Jordan, M. S. *Annu. Rev. Immunol.* **27**, 591–619 (2009).
4. Zhou, H. & Clapham, D. E. *Proc. Natl Acad. Sci. USA* **106**, 15750–15755 (2009).
5. Rijkers, G. T. & Griffioen, A. W. *Biochem. J.* **289**, 373–377 (1993).

## SOLAR PHYSICS

# Waves galore

**Wave energy has long been proposed to be a source of the hot solar corona and fast solar wind. Direct measurements made by spacecraft have finally established that coronal waves are ubiquitous and can have the required energy. SEE LETTER P.477**

PETER CARGILL & INEKE DE MOORTEL

The Sun's outer atmosphere, the corona, is known to have a temperature in excess of 1 million kelvin — hotter than the visible surface by a factor of more than 200. The corona also expands into interplanetary space as a high-speed flow, the solar wind, which is supersonic, reaching speeds of several hundred kilometres per second near Earth. The existence of the corona and wind requires *in situ* deposition of energy, with the likely source being the coronal magnetic field. It has long been conjectured that a class of magnetic waves called Alfvén waves can contain the required energy, but direct detection has proved challenging. On page 477 of this issue, McIntosh *et al.*<sup>1</sup> identify Alfvén waves in the corona in such profusion that their role needs urgent reassessment.

On the basis of measurements of coronal radiation at extreme ultraviolet (EUV) and X-ray wavelengths, sustenance of the corona requires an energy input of  $10^2$ – $10^4$  W m<sup>-2</sup>, depending on location<sup>2</sup>. Alfvén waves are named after their proposer, the Nobel laureate Hannes Alfvén, and are analogous to waves on a string, with the string tension being replaced by a tension force associated with bent magnetic field lines, and the string mass density being replaced by the density of the ionized coronal gas (or plasma). They have well known properties<sup>3</sup>, are thought to be generated in the cool low solar atmosphere, to propagate into the corona, and have been directly measured by spacecraft in the interplanetary medium<sup>4</sup>. Direct detection of waves in the corona itself has been challenging, because it requires very high time-cadence EUV observations with high spatial resolution. However, sporadic detections have been made at high altitudes off the solar limb<sup>5</sup>.

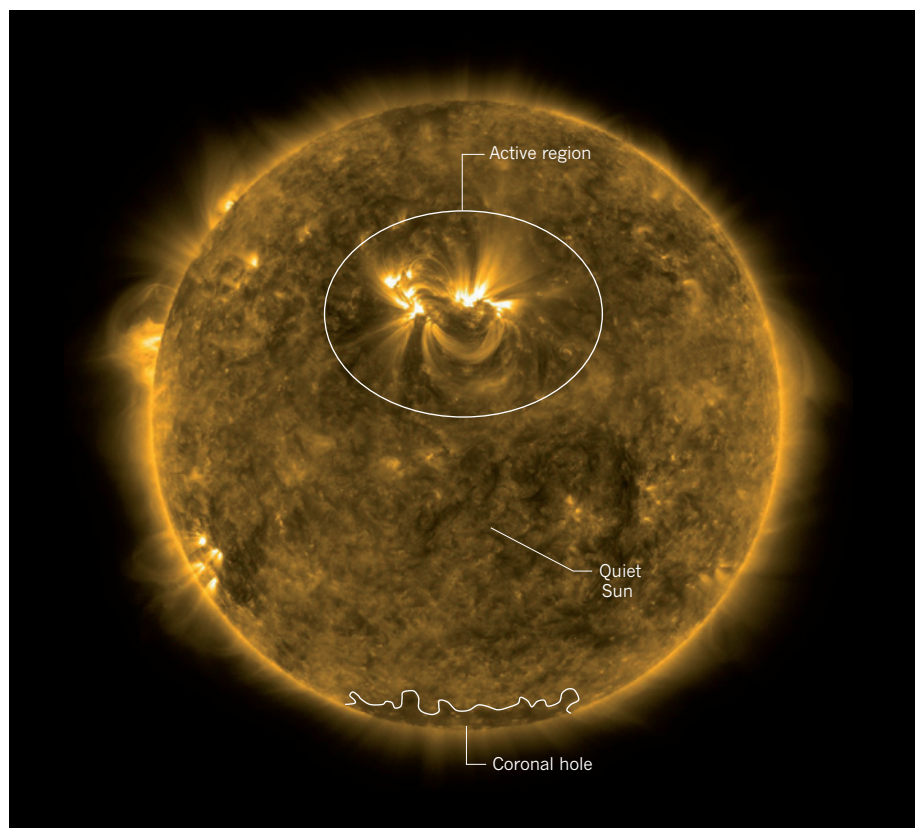
McIntosh *et al.*<sup>1</sup> present results that, for the first time, detect coronal Alfvén waves in such

profusion that they must now be regarded as ubiquitous. The key to this fundamental result is the remarkable time cadence, spatial resolution and sensitivity possible with the new Atmospheric Imaging Assembly on NASA's Solar Dynamics Observatory (SDO). EUV

images of structures and motions on the Sun can be generated, down to 1,000-km resolution, every 12 seconds, and this is fast and fine-scale enough to identify the presence of small-amplitude oscillatory displacements of structures throughout the corona. These oscillatory displacements have been interpreted as 'Alfvénic' waves on the basis of the large-scale (bulk) transverse oscillations of coronal structures that have been observed. (The term Alfvénic, rather than Alfvén, is used because the original definition by Alfvén applies only to oscillations in a uniform medium, which the highly structured corona is not.)

The second fundamental result concerns the estimate of the wave energy. McIntosh and colleagues divide the corona into three regions (Fig. 1). The first includes areas, known as coronal holes, in which the solar wind is believed to originate. These are associated with a magnetic field whose lines of force close in interplanetary space. The second comprises regions that are faint when viewed in EUV — known as the quiet Sun — but in which the lines of force close below the solar surface (Fig. 2). The third includes regions that are magnetically like the quiet Sun, but are sources of intense emission and are known as active regions.

In the coronal hole and quiet Sun regions,



**Figure 1 | The solar corona as seen in emission at 171 Ångströms.** This image, taken from the Atmospheric Imaging Assembly (AIA) instrument on the Solar Dynamics Observatory (SDO), shows coronal plasma at a temperature of about 1 million kelvin. McIntosh and colleagues<sup>1</sup> analyse three different types of region: coronal holes (source of the solar wind); quiet Sun; and active regions (sites of intense emission). In the coronal holes and quiet Sun, the power measured in magnetic oscillations known as Alfvén waves is adequate to account for the observed radiation.

AIA/SDO TEAM



## 50 Years Ago

### The First Anthocyanins appearing during the Ripening of Blueberries

We decided to investigate the blueberry, *Vaccinium myrtillus* L. ... It is known that the ripe blueberry contains at least seven anthocyanins ... Collecting the pigment from raw, reddish berries in sufficient amounts for identification was found to be too laborious a task. The pigments separated from raw berries were therefore compared with pigment fractions from ripe blueberries ... The two anthocyanins initially appearing during the maturation process of blueberries are both found to be cyanidin derivatives.

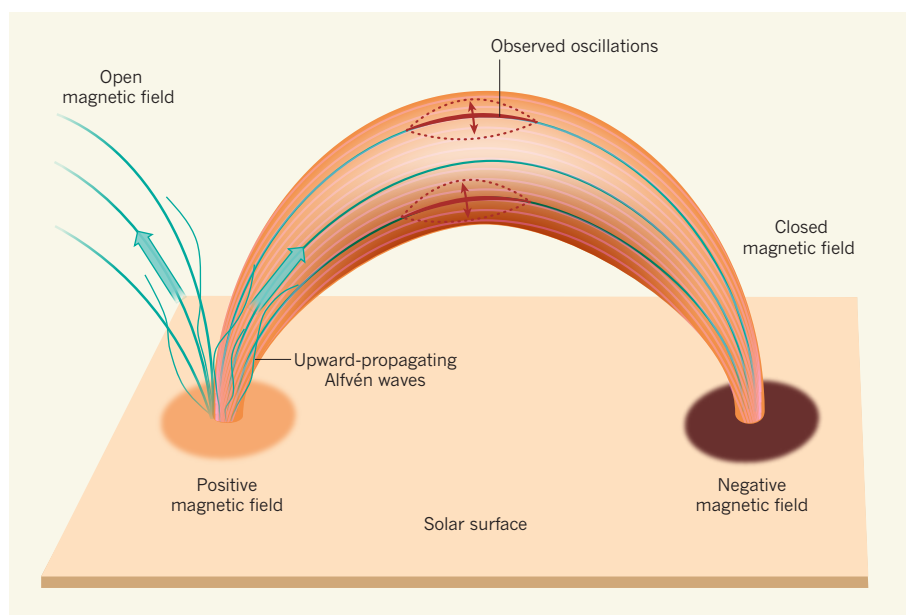
From *Nature* 29 July 1961

## 100 Years Ago

*The Law of Sex Determination and its Practical Application.* By Laura A. Calhoun (Mrs. E. E. Calhoun). Pp. 254. (New York: The Eugenics Publishing Co., 1910)

The theory suggested in this book is that "the sex of the embryo in man and the higher animals is determined in the ovary from which the ovum in question is developed." We shall not give away the ingenious author's practical recipe, but the general theory is that the right ovary is responsible for the males. This will be good news for those who believe that men are always in the right. ... The experiments ... reported in the *Journal of Genetics*, November, 1910, show that "in the rat it is not true that ova determining one sex are produced from one ovary, and those determining the opposite sex from the other, for each rat, with one ovary completely removed, produced young of both sexes." ... It is a well-intentioned book, but it does not contribute much to the difficult problem discussed.

From *Nature* 27 July 1911



**Figure 2 | Alfvén waves in two coronal geometries.** The solid lines represent magnetic field lines. On the left, the open magnetic field lines extend into interplanetary space, closing at a large distance from the Sun, and are representative of coronal holes. On the right, the field lines create closed structures in the corona, as occurs in the quiet Sun and active regions, with the lines closing beneath the surface. In both cases, the wavy blue lines represent upward-propagating transverse displacements (Alfvén waves), with the solid blue arrows showing the direction of energy transport. The dashed lines with red arrows represent the oscillatory displacements observed by McIntosh *et al.*<sup>1</sup> in active-region loops, interpreted as Alfvén waves.

the wave power estimated by McIntosh *et al.* is adequate to account for the coronal energy losses. In the observed active region, the measured power is inadequate by some margin. However, superposition of random motions along the line of sight (between the observer/spacecraft and the observed coronal structures) can lead to a large underestimation of the wave power (I. D. M. and D. J. Pascoe, unpublished work), so the authors' results can be considered as minimum values. In addition, higher time resolution may reveal more power at higher frequencies than currently observed.

Where does this leave things? A significant difficulty for coronal-heating mechanisms based on Alfvén waves is the weak damping of the waves, which can be likened to trying to damp the motion of a pendulum in a near-vacuum. Efficient damping is required to convert the kinetic and magnetic energy in the wave into heat. Idealized theories have suggested ways of enhancing the damping through the creation of strong, localized currents<sup>6</sup>, but one can see in movies from the SDO and the Japanese Hinode spacecraft that the solar corona is highly dynamic and structured, factors not accounted for in the basic theoretical models. The results of McIntosh *et al.* must lead to a thorough reassessment of the theory of waves in the solar atmosphere, with a focus on the (computational) modelling of complex, dynamic structures.

Finally, the argument about the origin of the corona has been framed for decades as a competition between the dissipation of small-scale

currents giving rise to impulsive heating<sup>7,8</sup>, direct plasma injection<sup>9</sup> and the dissipation of wave energy, with the first becoming more favoured. The new detection of 'waves galore' by McIntosh *et al.* suggests that this argument needs to be reconsidered as a matter of urgency. We suggest that the dynamic complexity revealed by Hinode and the SDO indicates that both processes are almost certainly at work, and hence a more pragmatic approach is called for to assess the relative contribution of each process in different regions. ■

**Peter Cargill** is at the *Blackett Laboratory, Imperial College, London SW7 2BZ, UK, and at the School of Mathematics & Statistics, University of St Andrews, Fife KY16 9SS, UK.* **Ineke De Moortel** is at the *School of Mathematics & Statistics, University of St Andrews.*

e-mails: [p.cargill@imperial.ac.uk](mailto:p.cargill@imperial.ac.uk); [ineke@mcs.st-and.ac.uk](mailto:ineke@mcs.st-and.ac.uk)

- McIntosh, S. W. *et al.* *Nature* **475**, 477–480 (2011).
- Withbroe, G. L. & Noyes, R. W. *Annu. Rev. Astron. Astrophys.* **15**, 363–387 (1977).
- Priest, E. R. *Solar Magnetohydrodynamics* (Reidel, 1982).
- Goldstein, M. L. & Roberts, D. A. *Phys. Plasmas* **6**, 4154–4160 (1999).
- Tomczyk, S. *et al.* *Science* **317**, 1192–1196 (2007).
- Goossens, M., Erdélyi, R. & Ruderman, M. S. *Space Sci. Rev.* (in the press).
- Parker, E. N. *Astrophys. J.* **330**, 474–489 (1988).
- Cargill, P. J. *Astrophys. J.* **422**, 381–393 (1994).
- De Pontieu, B. *et al.* *Science* **331**, 55–58 (2011).



# An *Archaeopteryx*-like theropod from China and the origin of Avialae

Xing Xu<sup>1,2</sup>, Hailu You<sup>3</sup>, Kai Du<sup>4</sup> & Fenglu Han<sup>2</sup>

*Archaeopteryx* is widely accepted as being the most basal bird, and accordingly it is regarded as central to understanding avialan origins; however, recent discoveries of derived maniraptorans have weakened the avialan status of *Archaeopteryx*. Here we report a new *Archaeopteryx*-like theropod from China. This find further demonstrates that many features formerly regarded as being diagnostic of Avialae, including long and robust forelimbs, actually characterize the more inclusive group Paraves (composed of the avialans and the deinonychosaurs). Notably, adding the new taxon into a comprehensive phylogenetic analysis shifts *Archaeopteryx* to the Deinonychosauria. Despite only tentative statistical support, this result challenges the centrality of *Archaeopteryx* in the transition to birds. If this new phylogenetic hypothesis can be confirmed by further investigation, current assumptions regarding the avialan ancestral condition will need to be re-evaluated.

The origin of the Avialae (defined as the most-inclusive clade containing *Passer domesticus* but not *Dromaeosaurus albertensis* or *Troodon formosus*; see Supplementary Information) represents one of the most heavily researched topics in evolutionary biology<sup>1,2</sup>. Being widely accepted as the most basal bird, *Archaeopteryx* has remained central to our understanding of avialan origins<sup>1,3</sup>. However, several recently reported basal avialans differ considerably from *Archaeopteryx* and instead share some salient similarities with oviraptorosaurs and, to a lesser degree, therizinosaurs<sup>4</sup>; conversely, *Archaeopteryx* and some *Archaeopteryx*-like theropods<sup>5</sup>, including the one reported here (Figs 1 and 2), possess some deinonychosaurian synapomorphies<sup>6</sup> (Fig. 3). These observations necessitate a re-evaluation of widely accepted hypotheses of paravian phylogeny; such an exercise will have significant implications for our understanding of avialan origins and related issues such as the origin of flight.

Theropoda Marsh, 1881  
Coelurosauria Huene, 1924  
Archaeopterygidae Huxley, 1871  
*Xiaotingia zhengi* gen. et sp. nov.

**Etymology.** The name is in honour of Zheng Xiaoting for his efforts in establishing the Shandong Tianyu Museum of Nature as a repository for vertebrate fossils from China.

**Holotype.** STM (Shandong Tianyu Museum of Nature) 27-2, an articulated skeleton, missing parts of the pelvis and hindlimbs and most of the caudal vertebrae, with some associated integumentary structures (Fig. 1; see Supplementary Information for the provenance and authenticity of the holotype specimen).

**Locality and horizon.** Linglongta, Jianchang, western Liaoning, China; Late Jurassic Tiaojishan Formation<sup>7</sup>.

**Diagnosis.** A small paravian with the following unique features: the maxillary posterior ramus has a depth at mid-length exceeding that of the dentary; the surangular has little lateral exposure and forms a wide, flat dorsal surface over the posterior part of the mandible; an extremely large surangular foramen extends over more than 6% of the total mandibular length; the posterior end of the mandible is blunt

and dorsoventrally expanded; the anteriormost caudal centra are less than half as long as the posterior dorsal centra; metacarpal IV is more robust than metacarpals II and III; and manual phalanx III-2 is longer than metacarpal III (we identify the three manual digits of *Xiaotingia* and other maniraptorans as II-III-IV, rather than as I-II-III as in many other studies<sup>8</sup>).

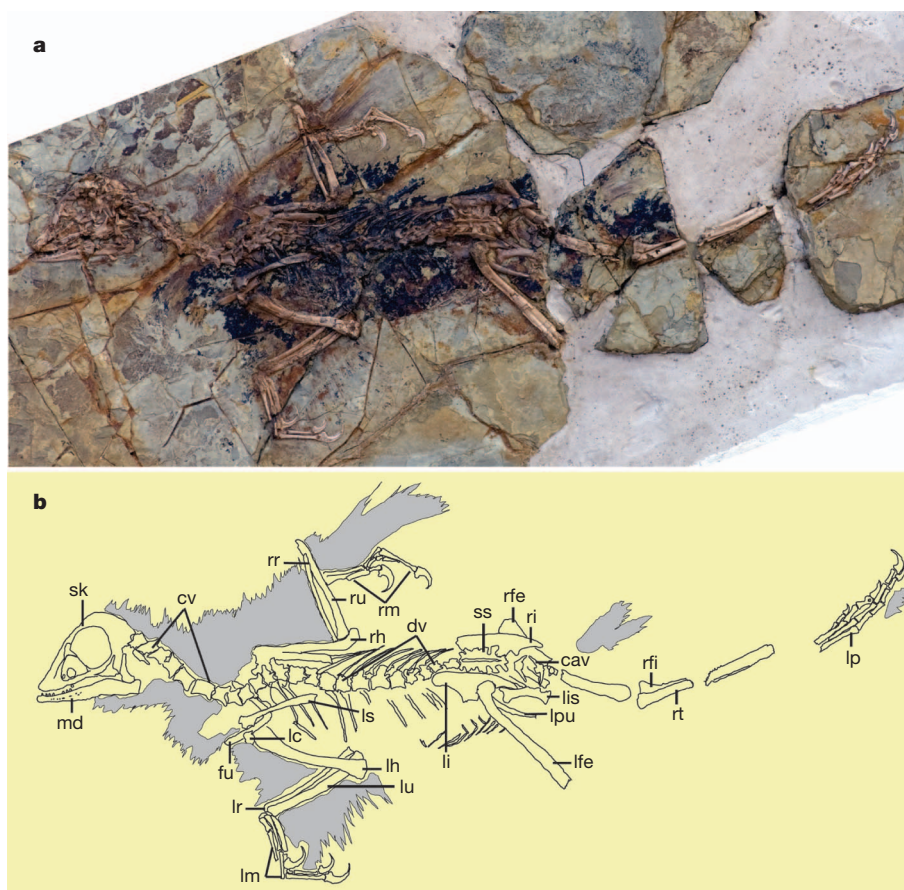
## Morphological description and comparison

The holotype specimen of *Xiaotingia zhengi* has completely closed neurocentral sutures on all exposed vertebrae and has a completely fused synsacrum, indicative of a late ontogenetic stage (probably adult). The body mass is estimated to be 0.82 kg based on an empirical bivariate equation<sup>9</sup>, similar to values calculated for other basal paravians<sup>10</sup>.

As in many maniraptorans including *Archaeopteryx*<sup>11</sup>, the antorbital fenestra is considerably shorter anteroposteriorly than it is high dorsoventrally. Similar to troodontids<sup>11</sup> and possibly *Archaeopteryx* (Fig. 3b), the descending process of the lacrimal is inset relative to the anterior and posterior processes (Fig. 2a). As in *Archaeopteryx*, *Anchiornis*, troodontids and some basal dromaeosaurids<sup>11</sup>, the jugal has a posterodorsally oriented, mediolaterally thick postorbital process and a small quadratojugal process that fails to extend as far posteriorly as the postorbital process (Fig. 2a). The posteriorly strongly curved quadrate bears a small pterygoid ramus, similar to the condition in *Archaeopteryx*, *Anchiornis* and some basal deinonychosaurs<sup>11</sup>. The pterygoid ramus is larger in basal avialans such as scansoriopterygids, *Sapeornis* and *Jeholornis*, and in oviraptorosaurs<sup>12</sup>. As in *Archaeopteryx* (Fig. 3c), *Anchiornis*, troodontids<sup>13</sup> and some basal dromaeosaurids<sup>14</sup>, the dentary bears a groove that widens posteriorly and contains a row of foramina (Fig. 2a). The retroarticular process is minimal, and the posterior end of the mandible is blunt as in confuciusornithids<sup>15</sup>.

The dentary tooth count (probably fewer than 10) is smaller than in other deinonychosaurs but greater than in basal avialans and oviraptorosaurs<sup>4</sup>. The teeth in the symphyseal region appear to be closely packed as in *Anchiornis*, troodontids and some basal dromaeosaurids<sup>11,13</sup>. The tooth crowns are short apicobasally and thick labiolingually, basally bulbous with a constriction below the bulbous

<sup>1</sup>College of Life Science, Linyi University, Shuangling Road, Linyi City, Shandong 276005, China. <sup>2</sup>Key Laboratory of Evolutionary Systematics of Vertebrates, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, 142 Xiwai Street, Beijing 100044, China. <sup>3</sup>Institute of Geology, Chinese Academy of Geological Sciences, 26 Baiwanzhuang Road, Beijing 100037, China. <sup>4</sup>Department of Biology, Capital Normal University, 105 Xisanhuan North Road, Beijing 100037, China.



**Figure 1** | *Xiaotingia zhengi* holotype (STM 27-2). **a**, **b**, Photograph (**a**) and line drawing (**b**). Integumentary structures in **b** are coloured grey. cav, caudal vertebra; cv, cervical vertebra; dv, dorsal vertebra; fu, furcula; lc, left coracoid; lfe, left femur; lh, left humerus; li, left ilium; lis, left ischium; lm, left manus; lp,

left pes; lpu, left pubis; lr, left radius; ls, left scapula; lu, left ulna; md, mandible; rfe, right femur; rfi, right fibula; rh, right humerus; ri, right ilium; rm, right manus; rr, right radius; rt, right tibiotarsus; ru, right ulna; sk, skull; ss, synsacrum.

swelling, and nearly symmetrical in labial view. They are similar in their general morphology to those of some basal avialans<sup>16,17</sup>.

The posterior cervical vertebrae have strongly divergent postzygapophyses, so that each vertebra is more than twice as wide as it is long. Pneumatic foramina are seen in the middle and posterior dorsal vertebrae (Fig. 2b), in contrast to the condition in most basal deinonychosaurs<sup>11</sup>. Five sacral vertebrae form a short synsacrum (less than 60% as long as the ilium), as in other archaeopterygids and basal deinonychosaurs. The zygapophyses of the sacral vertebrae are fused to form a platform lateral to the fused neural spines, a feature also known in dromaeosaurids and basal troodontids<sup>18</sup>. The anteriormost caudal centra are less than half as long as the posterior dorsal centra and have long, slender and distally tapering transverse processes (Fig. 2c), as in troodontids<sup>5,13</sup>.

The boomerang-shaped furcula is more robust than those of most other non-avian theropods, has an interclavicular angle of about 75°, and bears a small acromial process (Fig. 2b) as in *Anchiornis* and *Archaeopteryx* (Fig. 3d). The scapula has a strongly laterally everted acromial process overhanging a groove along the lateral surface and also bears a distinct short groove along the ventral edge immediately distal to the glenoid fossa, a feature also known in some previously described basal deinonychosaurs including *Anchiornis*. The coracoid has a relatively narrow proximal end and bears a fossa on the posterior surface (Fig. 2b), as in dromaeosaurids<sup>11</sup>.

The relatively long humerus is as robust as the femur. Metacarpal IV is the most robust metacarpal, and extends distally beyond metacarpal III (Fig. 2d), a feature known in enantiornithines. As in some basal oviraptorosaurs<sup>19</sup>, the phalangeal portion of the manus is much longer than the metacarpus. The penultimate manual phalanges are

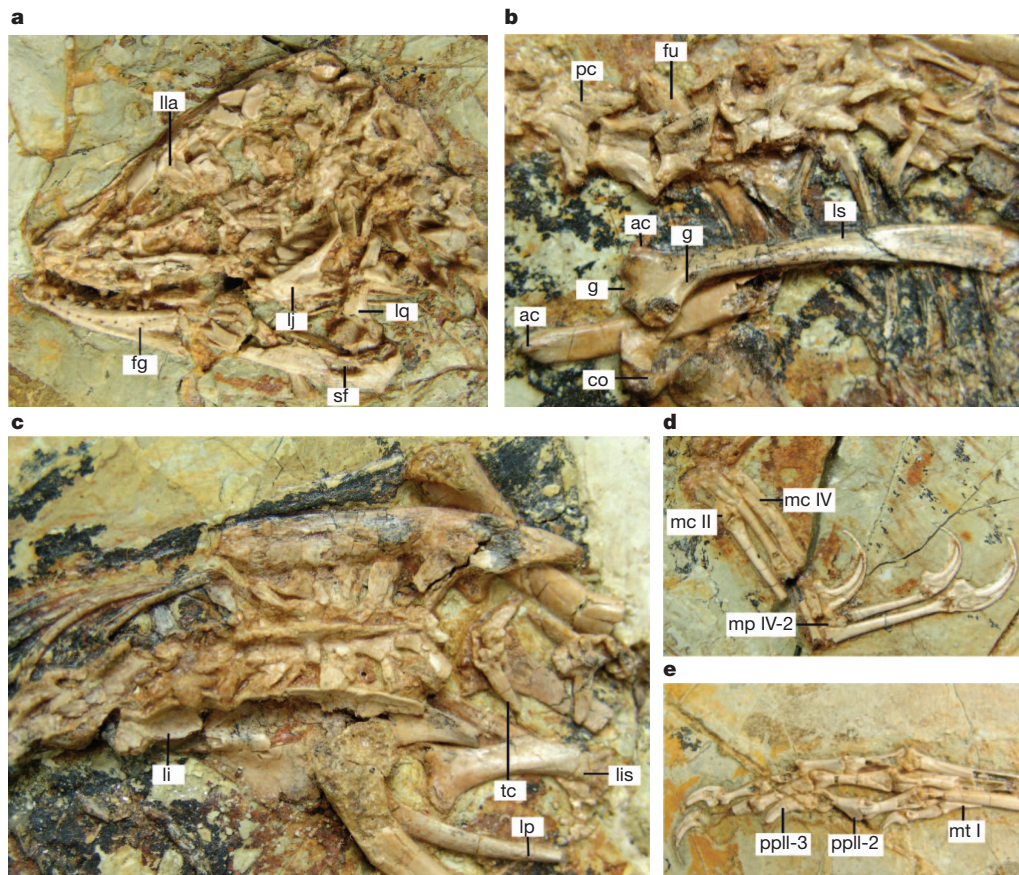
significantly longer than the more proximal ones, a feature indicative of arboreal habits<sup>20</sup>. Phalanx IV-1 is significantly longer than IV-2 and has a nearly immobile contact with the latter, as indicated by the absence of a pulley-like joint and the presence of a prominent proximoventral heel (Fig. 2d), as in *Archaeopteryx*<sup>16</sup> (Fig. 3e) and most dromaeosaurids<sup>11</sup>. The collateral ligament pits appear to be poorly developed.

The pre-acetabular process of the ilium is long (about 1.6 times as long as the postacetabular process) and anteriorly convex, as in other basal paravians<sup>11</sup>. The postacetabular process is rod-like and has a flat, thickened ventral surface as in *Archaeopteryx* and some basal troodontids<sup>11</sup>. The ischium has a groove along its anterior margin as in dromaeosaurids<sup>11</sup>.

The second pedal digit is similar to those of *Archaeopteryx*<sup>16</sup> (Fig. 3h), *Anchiornis*<sup>5</sup> and other deinonychosaurs in having a prominent dorsal expansion and a dorsally located lateral collateral ligament fossa at the distal ends of both phalanx II-1 and phalanx II-2 (Fig. 2e), indicating that the second digit was highly extensible. Also as in troodontids and dromaeosaurids, phalanx II-2 has a medially positioned ridge on the ventral surface near the proximal end. Unlike in troodontids and dromaeosaurids<sup>21,22</sup>, phalanx II-2 is not reduced in length and lacks a prominent proximoventral heel, and the ungual of digit II is only slightly longer than that of digit III and lacks a large flexor tubercle.

Faint feather impressions can be seen around the whole skeleton, including the skull, vertebral column, forelimbs and hindlimbs (Fig. 1). Some faint integumentary impressions are even preserved near the pedal phalanges, a feature also known in *Anchiornis*<sup>6</sup>. Unfortunately, the feathers are too poorly preserved for details of their





**Figure 2 | Selected skeletal elements of STM 27-2.** **a**, Skull and mandible. **b**, Middle presacral vertebrae, furcula and left scapulocoracoid. **c**, Pelvis and anterior caudal vertebrae. **d**, Left manus. **e**, Left pes. ac, acromial process; co, coracoid; fg, groove with foramina; fu, furcula; g, groove; li, left ilium; lis, left

ischium; lj, left jugal; lla, left lacrimal; lp, left pubis; lq, left quadrate; ls, left scapula; mc II, metacarpal II; mc IV, metacarpal IV; mp IV-2, manual phalanx IV-2; mt I, metatarsal I; pc, posterior cervical; ppII-2, pedal phalanx II-2; ppII-3, pedal phalanx II-3; sf, surangular foramen; tc, caudal transverse process.

structure to be apparent. The feathers near the femur are quite long, measuring more than 55 mm. The presence of such long femoral feathers is consistent with the tetrapterygian condition seen in several other basal paravian taxa<sup>6,23</sup>.

### Implications for paravian phylogeny

We have added *Xiaotingia* into a comprehensive phylogenetic analysis, which places *Xiaotingia* and *Anchiornis* within the Archaeopterygidae (Fig. 4 and Supplementary Information). Salient synapomorphies of the Archaeopterygidae include: manual phalanx III-1 more than twice as long as IV-1 (character state 292.1 in Supplementary Information); manual phalanx IV-3 markedly longer than IV-1 and IV-2 combined (character state 302.2 in Supplementary Information); furcula lateral end with L-shaped cross-section (character state 369.1 in Supplementary Information); and ventral notch between distal portion of obturator process and ischial shaft (character state 307.0 in Supplementary Information).

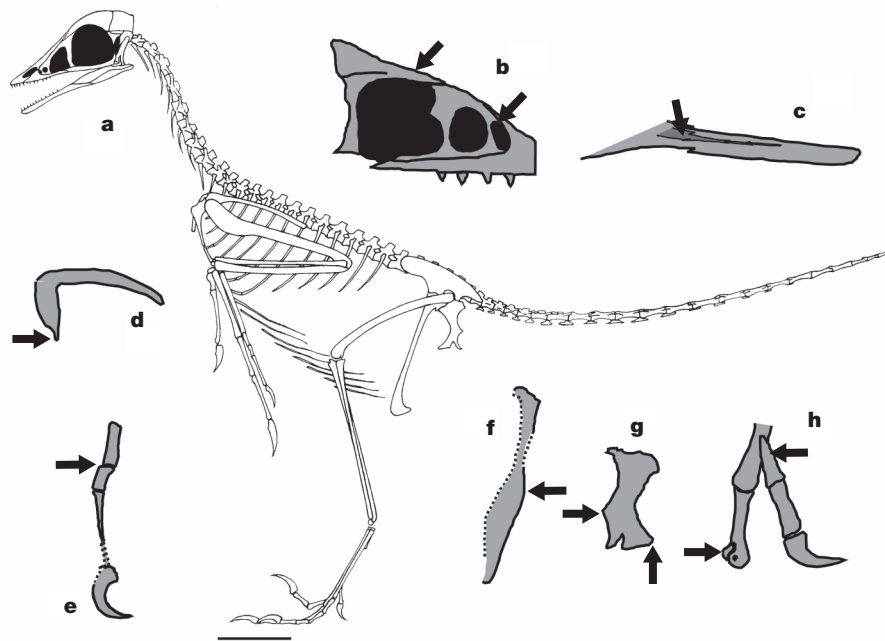
The most important result of our analysis is the removal of the Archaeopterygidae from the Avialae and its placement at the base of the Deinonychosauria, which challenges the long-held opinion that *Archaeopteryx* represents a pivotal taxon for understanding the transition to birds by virtue of having a phylogenetic position near the very base of the Avialae<sup>10,21,24–27</sup>. Derived features shared by *Archaeopteryx* and other deinonychosaurians include a large promaxillary fenestra (character state 363.1 in Supplementary Information), a T-shaped lacrimal with a long anterior process (character state 372.1 in Supplementary Information), a groove that widens posteriorly on the dentary (character state 72.1 in Supplementary Information), a manual phalanx IV-2 that is significantly shortened relative to IV-1

(character state 294.1 in Supplementary Information), a short ischium that bears a distally located obturator process as well as a posterodistal process (character states 171.2 and 334.1 in Supplementary Information), and a highly extensible pedal digit II (character state 323.1 in Supplementary Information), among others. Previous studies noted the striking similarities between *Archaeopteryx* and other deinonychosaurians<sup>16,28</sup>, and a close relationship between *Archaeopteryx* and dromaeosaurids has been proposed<sup>22</sup>, but to our knowledge we are the first to present a numerical phylogenetic analysis supporting deinonychosaurian affinities for the Archaeopterygidae.

It should be noted that our phylogenetic hypothesis is only weakly supported by the available data. Bremer support and bootstrap values for the recovered coelurosaurian subclades are, in general, low, and a bootstrap value less than 50% and a Bremer support value of 2 are obtained for a monophyletic Deinonychosauria including the Archaeopterygidae (see Supplementary Information). This low support is partly caused by various homoplasies, many of which are functionally significant, that are widely distributed across coelurosaurian phylogeny<sup>29</sup>. *Xiaotingia* possesses salient anatomical features also seen in different paravian taxa, further highlighting the phenomenon of widespread homoplasy. This phenomenon is also seen in some other major transitions, including the origins of major mammalian groups<sup>30</sup>, and creates difficulties in recovering robust phylogenies.

### Morphology and systematics of *Archaeopteryx*

Although *Archaeopteryx* has been known for about 150 years, debate continues regarding various aspects—including even some skeletal morphological features—of this extremely important taxon<sup>16,31</sup> (Fig. 3a). Recent findings, particularly the discovery of the tenth



**Figure 3 | Selected skeletal elements of *Archaeopteryx*.** **a**, Skeletal reconstruction (modified from Fig. 6.53 in ref. 16). **b**, Preorbital region of the Thermopolis specimen in lateral view. Arrows point to the long anterior process of the lacrimal and the large promaxillary fenestra. **c**, Anterior half of the mandible of the Eichstätt specimen in lateral view. Arrow points to a posteriorly widening groove. **d**, The furcula of the London specimen in oblique view. Arrow points to an L-shaped cross-section of the lateral end of the furcula. **e**, Manual digit IV of the Berlin specimen in dorsal view. Arrow points to the rigid connection between the long phalanx IV-1 and the short IV-2. **f**, Right

specimen, have greatly improved our knowledge of the morphology of *Archaeopteryx*<sup>16,28,31</sup>. In addition to the similarities between *Archaeopteryx*, *Xiaotingia* and some other deinonychosaurs described above, we provide further information to highlight the similarities between *Archaeopteryx*, *Anchiornis*, *Xiaotingia* and other deinonychosaurs on the one hand, and the differences between *Archaeopteryx* and other widely accepted basal avialans on the other.

The skull of *Archaeopteryx* is, in general, similar to those of *Anchiornis*, *Xiaotingia* and other deinonychosaurs in having a subtriangular lateral profile produced by a shallow snout and expanded postorbital region<sup>16</sup> (Fig. 4). In most basal avialans, including *Epidexipteryx*, *Sapeornis* and *Jeholornis*<sup>4,32</sup>, the skull is relatively tall and short with a deep, short snout, more reminiscent of the oviraptorosaurian condition (Fig. 4). As in *Anchiornis*, *Xiaotingia* and other basal deinonychosaurs<sup>6,11,16</sup>, the orbit is proportionally large and the infratemporal fenestra is extremely narrow anteroposteriorly and strongly inclined posteriorly. For comparison, oviraptorosaurs and basal avialans have a proportionally smaller orbit and a larger infratemporal fenestra that is much wider anteroposteriorly and less posteriorly inclined<sup>4,15,32</sup>. The external naris is ventrally located as in *Anchiornis*, *Xiaotingia* and other basal deinonychosaurs, in contrast to the high naris of oviraptorosaurs and basal avialans such as *Epidexipteryx* and *Jeholornis*<sup>4,32</sup>.

The premaxilla of *Archaeopteryx* is shallow in lateral view and much smaller than the maxilla, as in many theropods including deinonychosaurs<sup>11</sup>. In oviraptorosaurs and basal avialans such as *Epidexipteryx*, *Sapeornis* and *Jeholornis*<sup>4,32</sup>, the premaxilla is deep, and larger than the maxilla. The anteroposterior length of the antorbital fossa considerably exceeds its dorsoventral height, as in most theropods including *Anchiornis*, *Xiaotingia* and deinonychosaurs<sup>6,13,18</sup>. In oviraptorosaurs and basal avialans<sup>12,32</sup>, the opposite is true, and the antorbital fenestra within the fossa is thus much higher than anteroposteriorly long in lateral view. The promaxillary fenestra is large (Fig. 3b) as in *Anchiornis*, *Xiaotingia* and basal deinonychosaurs<sup>6,11,16,33</sup>—the

pubis of the Solnhofen specimen in posterior view. Arrow points to a lateral expansion at the pubic mid-shaft. **g**, Right ischium of the Thermopolis specimen in lateral view. Arrows point to the distally located obturator process and a triangular posterodistal process. **h**, Right pedal digits I and II of the Thermopolis specimen in oblique view. Arrow points to the medially positioned pedal digit I and the prominent dorsal expansion at the distal end of phalanx II-1. Most of the illustrated features here are only seen in archaeopterygids and other Deinonychosauria. Scale bar: 3 cm (**a**).

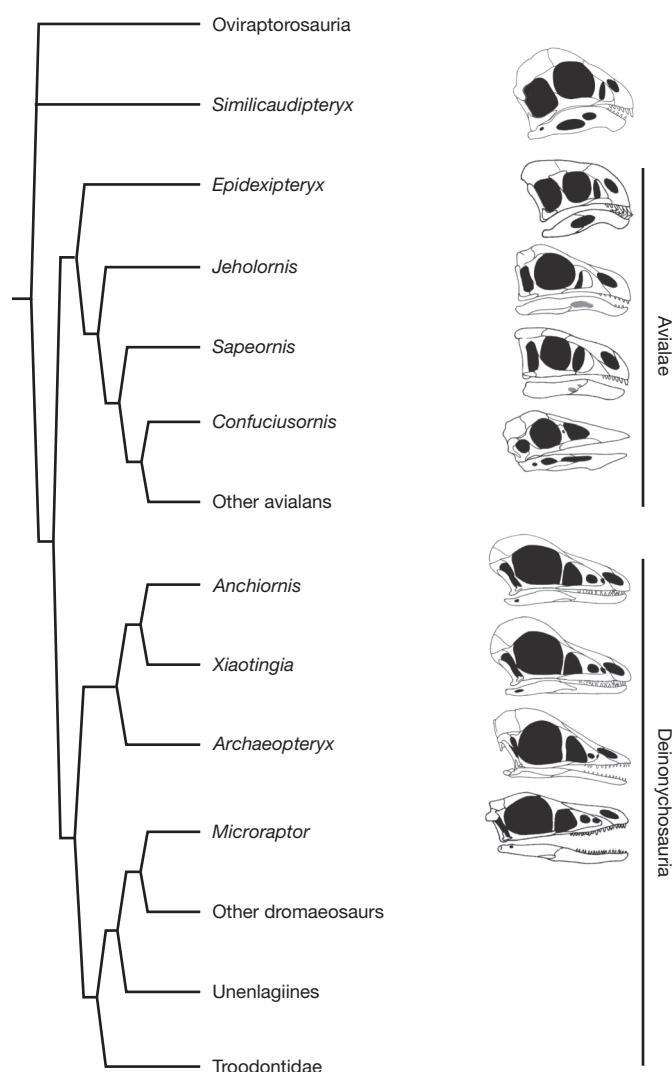
promaxillary fenestra, if present, is very small in other non-avian theropods. Many other theropods, including oviraptorosaurs and basal avialans such as *Epidexipteryx*, *Sapeornis* and *Jeholornis*<sup>4,12</sup>, lack a promaxillary fenestra (Fig. 4). The lacrimal has a long anterior process, close in length to the descending process and extending anteriorly to a point close to the anterior border of the antorbital fenestra, a feature also seen in deinonychosaurs<sup>11,13,18</sup>. In most other theropods and particularly in oviraptorosaurs and other basal avialans, the anterior process is proportionally much shorter. The lacrimal also has a posterior process, albeit a small one, as in oviraptorosaurs, *Anchiornis*, *Xiaotingia*, dromaeosaurids and troodontids, and the process is directed almost straight posteriorly as in *Anchiornis*, *Xiaotingia* and deinonychosaurs<sup>13,18</sup>. In oviraptorosaurs, the posterior process points posterodorsally, which seems also to be the case in some basal avialans<sup>15</sup>.

The mandible of *Archaeopteryx* is long and slender as in *Anchiornis*, *Xiaotingia* and basal deinonychosaurs<sup>6,11</sup>. For comparison, basal avialans all have oviraptorosaur-like mandibles: the mandible is relatively robust, the external mandibular fenestra is large and anteriorly located, and the dentary has a convex dorsal margin and a concave ventral one (however, the external mandibular fenestra is poorly known in *Jeholornis* and *Sapeornis*)<sup>4,32</sup>.

As in *Anchiornis* and basal deinonychosaurs<sup>11,34</sup>, the dorsal vertebrae of *Archaeopteryx* bear no distinct pneumatic foramina and instead have shallow, elongate depressions on the lateral surface of the centrum. In oviraptorosaurs and basal avialans such as *Jeholornis* and *Sapeornis*, the dorsal vertebrae bear distinct pneumatic foramina<sup>35</sup>. *Archaeopteryx* has five sacral vertebrae as in *Anchiornis*, *Xiaotingia*, basal troodontids and basal dromaeosaurids<sup>11,33</sup>. By contrast, basal avialans have a greater number of sacral vertebrae<sup>35</sup>.

The scapula is significantly shorter and more slender than the humerus, a feature also seen in other paravians<sup>11,35</sup>. Similar to the condition in other deinonychosaurs, the coracoid bears a distinct





**Figure 4 | A simplified cladogram showing the systematic position of *Xiaotingia* among the Coelurosauria (see Supplementary Information).** Morphological features in grey areas need confirmation by better preserved specimens. Taxa recovered as basal avialans by our analysis are more similar in general morphology to the oviraptorosaurs than to the archaeopterygids and basal deinonychosaurs.

subglenoid fossa. An ossified sternum and uncinat processes are absent as in *Anchiornis*, *Xiaotingia* and troodontids. The humerus has a proximodistally long internal tuberosity, as in *Anchiornis*, *Xiaotingia* and other deinonychosaurs<sup>11</sup>. The length of manual phalanx IV-3 is considerably greater than the combined lengths of IV-1 and IV-2, a feature also seen in *Anchiornis*, *Xiaotingia* and other deinonychosaurs.

As in *Anchiornis*, *Xiaotingia* and deinonychosaurs<sup>14,36,37</sup>, but unlike in basal avialans<sup>17</sup>, the pre-acetabular process of the ilium is relatively deep. The supraacetabular crest is distinct as in *Xiaotingia*, *Anchiornis* and basal deinonychosaurs. In oviraptorosaurs and basal avialans the supraacetabular crest is absent. In the Solnhofen specimen, a lateral expansion is present on the mid-shaft of the pubis (Fig. 3f), a feature also seen in basal dromaeosaurids and troodontids<sup>11</sup>. The very short and wide ischium has a distally located obturator process (Fig. 3g), as in *Anchiornis* and basal deinonychosaurs<sup>11,33,37</sup>. In most maniraptorans including oviraptorosaurs, the ischium is short, but not to the degree seen in *Archaeopteryx* and deinonychosaurs, and in all basal avialans the ischium has a different shape: relatively long and slender, posteriorly curved, and without an obturator process. The ischium has a distally located process on the posterior margin (Fig. 3g) as in basal

deinonychosaurs and *Xiaotingia*, although the posterior margin of the ischium of *Jeholornis* admittedly seems to bear a large convexity. A trait uniquely shared with *Anchiornis* (condition unknown in *Xiaotingia*) is the constricted base of the distally located obturator process.

The metatarsus of *Archaeopteryx* approaches the arctometatarsalian condition<sup>28</sup> in that the proximal end of the third metatarsal is laterally compressed as in *Anchiornis* and basal deinonychosaurs. In basal avialans, metatarsal III is not laterally compressed<sup>4,17</sup>.

As described above, *Archaeopteryx* is more similar to *Anchiornis*, *Xiaotingia* and basal deinonychosaurs than to known basal avialans and oviraptorosaurs in numerous features, some of which are uniquely shared. On the other hand, basal avialans such as scansoriopterygids, *Sapeornis*, *Jeholornis* and the confuciusornithids are more similar to oviraptorosaurs than to *Archaeopteryx*, *Anchiornis*, *Xiaotingia* and basal deinonychosaurs in many features, particularly cranial and vertebral ones. This supports the hypothesis that *Archaeopteryx*, *Anchiornis* and *Xiaotingia* are referable to the Deinonychosauria, a hypothesis consistent with some previous work on *Anchiornis*<sup>6,38</sup>.

Although *Archaeopteryx* is placed within the Avialae by nearly all numerical phylogenetic studies<sup>10,11,21,24–26,39,40</sup>, some recent studies have demonstrated that some of the suggested synapomorphies purportedly shared by *Archaeopteryx* and basal avialans are questionable. For example, two salient avialan features—the absence of a jugal process on the palatine and the presence of a reversed hallux—are now considered to be absent in *Archaeopteryx*<sup>28,31</sup> (Fig. 3h). Some other suggested synapomorphies are present in recently described basal deinonychosaurs, and are thus likely to represent paravian rather than avialan synapomorphies<sup>23,37</sup>. These features include an antorbital fossa that is dorsally bordered by the nasal and lacrimal, a relatively small number of caudal vertebrae, a relatively large proximodorsal process of the ischium, a relatively long pre-acetabular process of the ilium, and fusion of the proximal part of the metatarsus<sup>11,37,41</sup>. Consequently, there are few derived features shared by *Archaeopteryx* and basal avialans but absent in basal deinonychosaurs, thus documented morphological support for the avialan affinities of *Archaeopteryx* is fairly weak. The alternative hypothesis that *Archaeopteryx*, *Anchiornis* and *Xiaotingia* are all deinonychosaurs is better supported by the available morphological data, and these taxa share with some basal deinonychosaurs some unique features unknown in any other theropod group (Figs 3 and 4; see also Supplementary Information).

Within the Deinonychosauria, *Archaeopteryx* is more similar to *Anchiornis* and *Xiaotingia* than to dromaeosaurids and troodontids in many features, although few of these features are uniquely shared by the three taxa. Of note, however, are some unique features related to the pelvis. For example, the ischium appears to be proportionally even shorter in *Archaeopteryx* and *Anchiornis* than in other deinonychosaurs, and these two taxa also share a basally constricted obturator process (condition unknown for both characters in *Xiaotingia*). On the other hand, *Archaeopteryx*, *Anchiornis* and *Xiaotingia* lack many derived similarities shared by troodontids and dromaeosaurids, such as lateral exposure of the splenial, a muscle scar on the deltopectoral crest, and an enlarged, raptorial ungual on pedal digit II. This suggests that *Archaeopteryx*, *Anchiornis* and *Xiaotingia* are probably most closely related to each other, whereas dromaeosaurids and troodontids form a separate clade within the Deinonychosauria (see additional comparative figures in Supplementary Information).

### Implications for avialan origins

The discovery of *Xiaotingia* further demonstrates that many features previously regarded as distinctively avialan actually characterize the more inclusive Paraves. For example, proportionally long and robust forelimbs are optimized in our analysis as a primitive character state for the Paraves (see Supplementary Information). The significant lengthening and thickening of the forelimbs indicates a dramatic shift

in forelimb function at the base of the Paraves, which might be related to the appearance of a degree of aerodynamic capability. This hypothesis is consistent with the presence of flight feathers with asymmetrical vanes in both basal avialans and basal deinonychosaurs<sup>6,23</sup>.

All taxa recovered as basal avialans by our analysis, such as the scansoriopterygids, *Sapeornis* and *Jeholornis*, resemble oviraptorosaurs and to a lesser degree therizinosaurs<sup>4</sup> but differ from deinonychosaurs including archaeopterygids in having such cranial and dental characteristics as a dorsoventrally high premaxilla that is significantly larger than the maxilla, a dorsally positioned external naris, a dorsoventrally tall antorbital fossa, a jugal with a relatively vertical postorbital process and a long quadratojugal process, a quadrate with a large pterygoid ramus, a relatively long parietal, an anteriorly downturned and strongly dorsally convex mandible, a large external mandibular fenestra, and enlarged anterior teeth. Some of these features are optimized by our analysis as synapomorphies of a clade containing the Oviraptorosauria, the Therizinosauroidea, the Avialae and the Deinonychosauria, but are lost in the last group (see Supplementary Information). Some previous phylogenetic analyses have placed the Oviraptorosauria within the Avialae<sup>42</sup>, and a recent study suggests that the Oviraptorosauria and Scansoriopterygidae are sister taxa, forming a clade at the base of the Avialae<sup>38</sup>. However, our analysis indicates that placing the Oviraptorosauria outside the Paraves is much more parsimonious than placing it within the Avialae (see Supplementary Information). In either case, many oviraptorosaur-like features are plesiomorphic for the Avialae. These features contribute to forming a relatively tall and robust cranium, in contrast to the shallower and more gracile cranium seen in the Deinonychosauria. These results invite a re-evaluation of the ancestral condition for birds from the perspective of morphology, behaviour and ecology. Under the phylogenetic framework shown in Fig. 4, a robust skull and a herbivorous diet (which has been suggested to characterize the Maniraptoriformes<sup>43,44</sup>) probably represent ancestral traits that are retained in basal birds, and the Deinonychosauria is exceptional in having a more gracile skull and a carnivorous diet.

Received 16 November 2010; accepted 10 June 2011.

- Feduccia, A. *The Origin and Evolution of Birds* 2nd edn (Yale Univ. Press, 1999).
- Zhou, Z.-H. The origin and early evolution of birds: discoveries, disputes, and perspectives from fossil evidence. *Naturwissenschaften* **91**, 455–471 (2004).
- Witmer, L. M. in *Mesozoic Birds: Above the Heads of Dinosaurs* (eds Chiappe, L. M. & Witmer, L. M.) 3–30 (Univ. California Press, 2002).
- Zhang, F. C. *et al.* A bizarre Jurassic maniraptoran from China with elongate ribbon-like feathers. *Nature* **455**, 1105–1108 (2008).
- Xu, X. *et al.* A new feathered maniraptoran dinosaur fossil that fills a morphological gap in avian origin. *Chin. Sci. Bull.* **54**, 430–435 (2009).
- Hu, D. Y. *et al.* A pre-Archaeopteryx troodontid from China with long feathers on the metatarsus. *Nature* **461**, 640–643 (2009).
- Xu, K. *et al.* *Jurassic System in the North of China (VII): the Stratigraphic Region of Northeast China* (Petroleum Industry Press, Beijing, 2003).
- Tamura, K. *et al.* Embryological evidence identifies wing digits in birds as digits 1, 2, and 3. *Science* **331**, 753–757 (2011).
- Christiansen, P. & Fariña, R. A. Mass prediction in theropod dinosaurs. *Hist. Biol.* **16**, 85–92 (2004).
- Turner, A. H. *et al.* A basal dromaeosaurid and size evolution preceding avian flight. *Science* **317**, 1378–1381 (2007).
- Xu, X. *Deinonychosaurian Fossils from the Jehol Group of Western Liaoning and the Coelurosaurian Evolution*. PhD thesis, Chinese Academy of Sciences (2002).
- Osmólska, H., Currie, P. J. & Barsbold, R. in *The Dinosauria* 2nd edn (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 165–183 (Univ. California Press, 2004).
- Makovicky, P. J. & Norell, M. A. in *The Dinosauria* 2nd edn (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 184–195 (Univ. California Press, 2004).
- Makovicky, P. J., Apestegui, S. & Agnolín, F. L. The earliest dromaeosaurid theropod from South America. *Nature* **437**, 1007–1011 (2005).
- Chiappe, L. M. *et al.* Anatomy and systematics of the Confuciusornithidae (Theropoda: Aves) from the late Mesozoic of Northeastern China. *Bull. Am. Mus.* **242**, 1–89 (1999).
- Wellnhofer, P. *Archaeopteryx-Der urvogel von Solnhofen* (Friedrich Pfeil, 2008).
- Zhou, Z. H. & Zhang, F. C. Anatomy of the primitive bird *Sapeornis chaoyangensis* from the Early Cretaceous of Liaoning, China. *Can. J. Earth Sci.* **40**, 731–747 (2003).
- Norell, M. A. & Makovicky, P. J. in *The Dinosauria* 2nd edn (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 196–209 (Univ. California Press, 2004).
- Ji, Q., Currie, P. J., Norell, M. A. & Ji, S.-A. Two feathered dinosaurs from northeastern China. *Nature* **393**, 753–761 (1998).
- Fröbisch, J. & Reisz, R. R. The Late Permian herbivore *Suminia* and the early evolution of arboreality in terrestrial vertebrate ecosystems. *Proc. R. Soc. Lond. B* **276**, 3611–3618 (2009).
- Gauthier, J. Saurischian monophyly and the origin of birds. *Memoirs Cal. Acad. Sci.* **8**, 1–55 (1986).
- Paul, G. *Predatory Dinosaurs of the World* (Simon & Schuster, 1988).
- Xu, X. *et al.* Four-winged dinosaurs from China. *Nature* **421**, 335–340 (2003).
- Sereno, P. C. The evolution of dinosaurs. *Science* **284**, 2137–2147 (1999).
- Holtz, T. R., Jr. A new phylogeny of the carnivorous dinosaurs. *Gaia* **15**, 5–61 (2000).
- Norell, M. A., Clark, J. M. & Makovicky, P. J. in *New Perspectives on the Origin and Evolution of Birds* (eds Gauthier, J. & Gall, L. F.) 49–67 (Yale Univ. Press, 2001).
- Choiniere, J. N. *et al.* A basal alvarezsaurid theropod from the early Late Jurassic of Xinjiang, China. *Science* **327**, 571–574 (2010).
- Mayr, G. *et al.* The tenth skeletal specimen of *Archaeopteryx*. *Zool. J. Linn. Soc.* **149**, 97–116 (2007).
- Holtz, T. R. J. in *New Perspectives on the Origin and Early Evolution of Birds* (eds Gauthier, J. & Gall, L. F.) 99–124 (Peabody Museum of Natural History, 2001).
- Luo, Z. X. Transformation and diversification in early mammal evolution. *Nature* **450**, 1011–1019 (2007).
- Mayr, G., Pohl, B. & Peters, S. A well-preserved *Archaeopteryx* specimen with theropod features. *Science* **310**, 1483–1486 (2005).
- Hu, D.-Y. *et al.* A new sapeornithid bird from China and its implication for early avian evolution. *Acta Geol. Sin.* **84**, 472–482 (2010).
- Xu, X., Norell, M. A., Wang, X.-L., Makovicky, P. J. & Wu, X.-C. A basal troodontid from the Early Cretaceous of China. *Nature* **415**, 780–784 (2002).
- Hwang, S. H. *et al.* New specimens of *Microaptor zhaoianus* (Theropoda: Dromaeosauridae) from northeastern China. *Am. Mus. Novit.* **3381**, 1–44 (2002).
- Zheng, X. T. *The Origin of Birds* (Shandong Science and Technology Press, 2009).
- Novas, F. E. & Puerta, P. F. New evidence concerning avian origins from the Late Cretaceous of Patagonia. *Nature* **387**, 390–392 (1997).
- Forster, C. A. *et al.* The theropod ancestry of birds: new evidence from the Late Cretaceous of Madagascar. *Science* **279**, 1915–1919 (1998).
- Xu, X., Ma, Q.-Y. & Hu, D.-Y. Pre-Archaeopteryx coelurosaurian dinosaurs and their implications for understanding avian origins. *Chin. Sci. Bull.* **55**, 3971–3977 (2010).
- Rauhut, O. W. M. The interrelationships and evolution of basal theropod dinosaurs. *Palaeontology* **69**, 1–215 (2003).
- Senter, P. A new look at the phylogeny of Coelurosauria (Dinosauria: Theropoda). *J. Syst. Palaeontology* **5**, 1–35 (2007).
- Novas, F. E. in *Feathered Dragons, Studies on the Transition From Dinosaurs to Birds* (eds Currie, P. J., Koppelhaas, E. B., Shugar, M. A. & Wright, J. L.) 150–168 (Indiana Univ. Press, 2004).
- Maryanska, T., Osmólska, H. & Wolsan, M. Avialan status for Oviraptorosauria. *Acta Palaeontol. Pol.* **47**, 97–116 (2002).
- Zanno, L. E. *et al.* A new North American therizinosaurid and the role of herbivory in “predatory” dinosaur evolution. *Proc. R. Soc. Lond. B* **276**, 3505–3511 (2009).
- Zanno, L. E. & Makovicky, P. J. Herbivorous ecomorphology and specialization patterns in theropod dinosaur evolution. *Proc. Natl Acad. Sci. USA* **108**, 232–237 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The authors thank X. Zheng for permission to study the holotype specimen of *Xiaotingia zhengi* and for discussions; P. Chen, L. Hou and Z. Zhou for coordinating the project; C. Sullivan for editing and commenting on the manuscript; R. Li for illustrations; T. Yu for preparing the specimen; X. Ding for editing the illustrations; D. Pol for help in using the TNT software package; P. Barrett, O. Rauhut, M. Kölbl-Ebert and M. Röper for facilitating access to the *Archaeopteryx* specimens under their care; and G. Mayr for discussions. This work was supported by grants from the National Natural Science Foundation of China, Chinese Academy of Sciences, and Special Funds for Major State Basic Research Projects of China.

**Author Contributions** X.X. designed the project, X.X., H.Y., K.D. and F.H. performed the research, and X.X. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to X.X. ([xingxu@vip.sina.com](mailto:xingxu@vip.sina.com)).



# Second messenger role for $Mg^{2+}$ revealed by human T-cell immunodeficiency

Feng-Yen Li<sup>1,2\*</sup>, Benjamin Chaigne-Delalande<sup>1\*</sup>, Chrysi Kanellopoulou<sup>1</sup>, Jeremiah C. Davis<sup>3</sup>, Helen F. Matthews<sup>1</sup>, Daniel C. Douek<sup>4</sup>, Jeffrey I. Cohen<sup>5</sup>, Gulbu Uzel<sup>6</sup>, Helen C. Su<sup>3</sup> & Michael J. Lenardo<sup>1</sup>

**The magnesium ion,  $Mg^{2+}$ , is essential for all life as a cofactor for ATP, polyphosphates such as DNA and RNA, and metabolic enzymes, but whether it plays a part in intracellular signalling (as  $Ca^{2+}$  does) is unknown. Here we identify mutations in the magnesium transporter gene, *MAGT1*, in a novel X-linked human immunodeficiency characterized by CD4 lymphopenia, severe chronic viral infections, and defective T-lymphocyte activation. We demonstrate that a rapid transient  $Mg^{2+}$  influx is induced by antigen receptor stimulation in normal T cells and by growth factor stimulation in non-lymphoid cells. *MAGT1* deficiency abrogates the  $Mg^{2+}$  influx, leading to impaired responses to antigen receptor engagement, including defective activation of phospholipase  $C\gamma 1$  and a markedly impaired  $Ca^{2+}$  influx in T cells but not B cells. These observations reveal a role for  $Mg^{2+}$  as an intracellular second messenger coupling cell-surface receptor activation to intracellular effectors and identify *MAGT1* as a possible target for novel therapeutics.**

$Mg^{2+}$  is the most abundant divalent cation in mammalian cells and is an essential cofactor for ATP, nucleic acids, and numerous enzymes in animals and plants<sup>1–3</sup>. Whether it serves as a second messenger in intracellular signalling is controversial<sup>4–8</sup>.  $Ca^{2+}$ , another alkaline earth metal, is well-established as a second messenger; this is because the concentration of free intracellular  $Ca^{2+}$ ,  $[Ca^{2+}]_i$ , is 0.1  $\mu M$  and that of free extracellular  $Ca^{2+}$ ,  $[Ca^{2+}]_o$ , is 1 mM, creating a steep gradient that favours voltage- and ligand-gated  $Ca^{2+}$  influx signals<sup>9</sup>. By contrast,  $[Mg^{2+}]_i$  is 10–30 mM and mostly complexed with ATP and other molecules. Only 1–5% (0.2–1 mM) is cytosolic free  $Mg^{2+}$ . However, this is 100-fold below its electrochemical equilibrium potential, which theoretically allows regulated  $Mg^{2+}$  influxes<sup>6,10</sup>.  $Mg^{2+}$  has been found to enhance lymphocyte activation in suboptimal  $Ca^{2+}$  concentrations by phytohaemagglutinin (PHA) but not ionomycin<sup>11–13</sup>. As ionomycin bypasses proximal T-cell antigen receptor (TCR) signals, optimal T-cell activation could require a magnesium-generated process upstream of  $Ca^{2+}$  signalling. Probes sensitive to  $Mg^{2+}$  have revealed changes in  $[Mg^{2+}]_i$  in lymphocytes following lectin stimulation<sup>14,15</sup>. Nevertheless, how extracellular  $Mg^{2+}$  promotes cellular activation signals is unknown.

The molecular elucidation of primary immunodeficiencies often yields novel insights into lymphocyte signal transduction<sup>16–18</sup>. TCR signalling is critical for T-lymphocyte selection during ontogeny and for peripheral responses against foreign pathogens<sup>19,20</sup>. Mutations in the zeta-chain-associated protein kinase of 70 kDa (ZAP70) tyrosine kinase in severe combined immunodeficiency patients illustrated its role in thymic development of CD8<sup>+</sup> T cells and peripheral T-cell activation<sup>19,21</sup>. Likewise, genetic defects in severe combined immunodeficiency patients in *ORAI1* revealed that it was a critical store-operated  $Ca^{2+}$  channel<sup>22</sup>. Idiopathic CD4 lymphocytopenia (ICL) is a rare immunodeficiency in which circulating CD4<sup>+</sup> T-cell counts are  $<300\text{ mm}^{-3}$  or  $<20\%$  of total T cells in the absence of HIV infection or other causes of lymphopenia<sup>23–25</sup>. The aetiology of ICL

is heterogeneous, with no infectious cause identified and a genetic aetiology suspected in some cases<sup>26–29</sup>. Here we have uncovered a second messenger role for  $Mg^{2+}$  in receptor-induced phospholipase  $C\gamma 1$  (PLC $\gamma 1$ ) activation and  $Ca^{2+}$  signalling by discovering an X-linked mutation in *MAGT1*, a highly selective transporter for  $Mg^{2+}$ , in a subset of ICL patients.

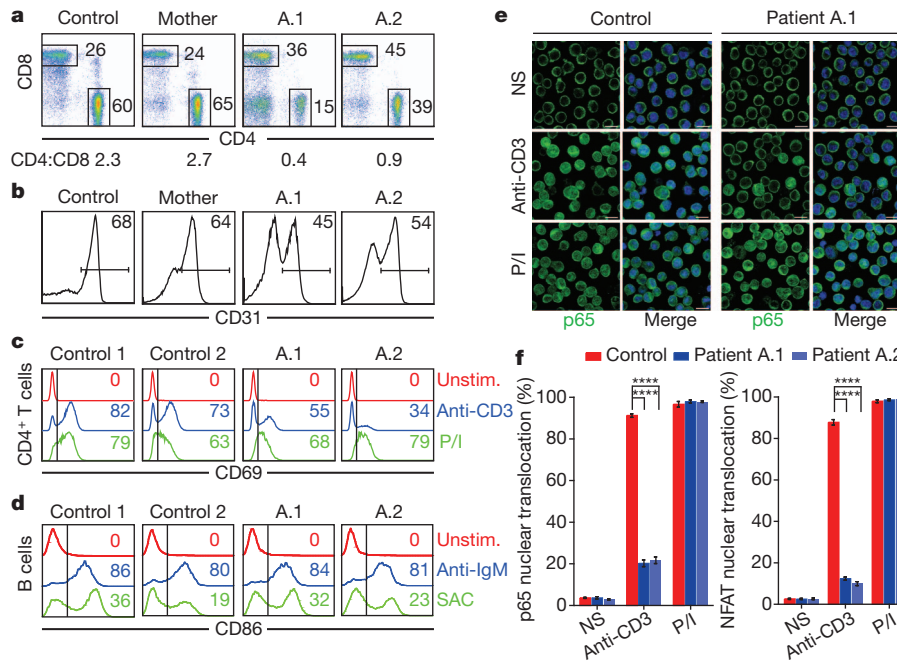
## Immune and gene defects in ICL patients

We examined two young brothers (patients A.1 and A.2) who exhibited recurrent infections, including chronic Epstein–Barr virus infection and low CD4<sup>+</sup> T-cell counts; other lymphocyte populations were normal or elevated, and immunoglobulin levels and vaccine responses were only intermittently defective (Supplementary Table 1, Supplementary Fig. 1a, b). Both patients had an inverted CD4:CD8 ratio and reduced CD31<sup>+</sup> cells in the naïve (CD27<sup>+</sup>, CD45RO<sup>–</sup>) CD4<sup>+</sup> T-cell population, suggesting diminished thymic output<sup>30–32</sup> (Fig. 1a, b). However, we also found pronounced defects in TCR-mediated activation events, including CD69, CD25, Fas (CD95) and CTLA-4 upregulation following OKT3 (agonistic anti-CD3) stimulation in peripheral blood mononuclear cells (PBMCs) (Fig. 1c, Supplementary Table 1, Supplementary Fig. 2a). Early TCR signalling events, such as NF- $\kappa$ B and NFAT nuclear translocation, were impaired (Fig. 1e, f). By contrast, the patients' T cells were fully activated by the downstream inducers phorbol 12-myristate 13-acetate (PMA) and ionomycin, implying a proximal TCR activation defect (Fig. 1c, e, f). The patients showed no defect in B-cell receptor (BCR) or toll-like receptor (TLR) stimulation of B cells (Fig. 1d, Supplementary Fig. 2b).

Given that family A harboured two affected boys (Fig. 2a), we tested for X-chromosome linkage by assessing lyonization, the process of X-chromosome inactivation by methylation in females<sup>33</sup>. Skewed lyonization reflects the reduced fitness of cells due to a deleterious X-chromosome mutation. We digested genomic DNA with the methylation-sensitive restriction endonuclease HpaII to eliminate

<sup>1</sup>Molecular Development Section, Lymphocyte Molecular Genetics Unit, Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>2</sup>Biomedical Sciences Graduate Program, University of California–San Francisco, San Francisco, California 94143, USA. <sup>3</sup>Human Immunological Diseases Unit, Laboratory of Host Defenses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>4</sup>Human Immunology Section, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>5</sup>Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>6</sup>Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

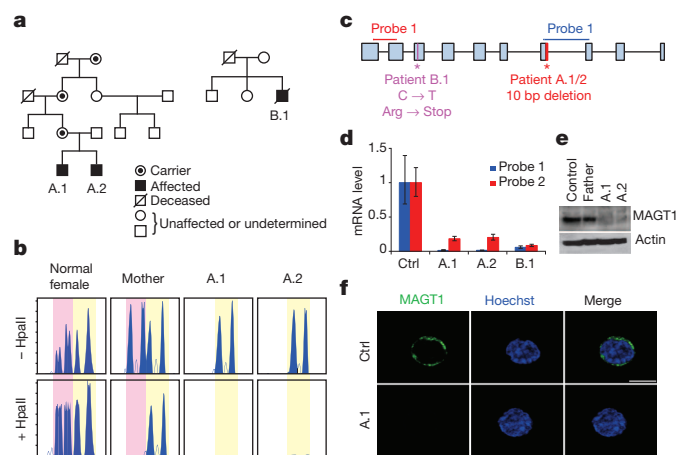
\*These authors contributed equally to this work.



**Figure 1 | Patients have a proximal TCR activation defect.** **a**, T-cell CD4 and CD8 expression and ratio. **b**, CD31 expression in naïve CD4<sup>+</sup>CD3<sup>+</sup> T cells. **c**, CD69 expression in CD4<sup>+</sup> T cells after 1  $\mu$ g ml<sup>-1</sup> anti-CD3 stimulation, PMA/ionomycin (P/I) stimulation or no stimulation (Unstim.). **d**, CD86 expression in purified B cells after stimulation with anti-IgM, SAC or no stimulation.

all active non-methylated X-chromosome DNA and analysed the remaining undigested (inactive) X chromosome by PCR at an indicator locus<sup>33</sup>. The mother of the two affected boys exhibited completely skewed lyonization with only the X chromosome inherited by her two sons inactivated in her T cells, strongly suggesting that she carries an X-linked genetic defect (Fig. 2b).

We therefore performed X-chromosome exon-capture and single-end next-generation sequencing on the mother and the two boys, yielding 18–20 million reads per subject with at least 10 $\times$  coverage



**Figure 2 | Patients have MAGT1-null mutations.** **a**, Pedigree of the families A (left) and B (right). **b**, X-chromosome inactivation assessed by digestion with (+HpaII) or without (–HpaII) methylation-sensitive enzyme HpaII. Peaks of PCR products from the different alleles are highlighted in yellow and pink. **c**, Schematic representation of the MAGT1 gene exons (boxes) and introns (lines), the mutations (\*) and the probes used for RT-PCR. **d**, RT-PCR showing expression of MAGT1 mRNA normalized to HPRT mRNA in T cells expressed as relative to normal control (CTRL). **e**, Expression of MAGT1 and actin control in T cells by immunoblot. **f**, Confocal images of T cells stained with anti-MAGT1 antibody (scale bars, 5  $\mu$ m).

**e**, Confocal imaging of p65 nuclear translocation after anti-CD3 or P/I stimulation or no stimulation (NS) (scale bar, 10  $\mu$ m). **f**, Percentage of cells with p65 (left) and NFAT (right) nuclear translocation. Numbers represent the percentage of cells in indicated gates. Error bars, s.e.m. ( $n = 3$ ); \*\*\*\* $P < 0.0001$ .

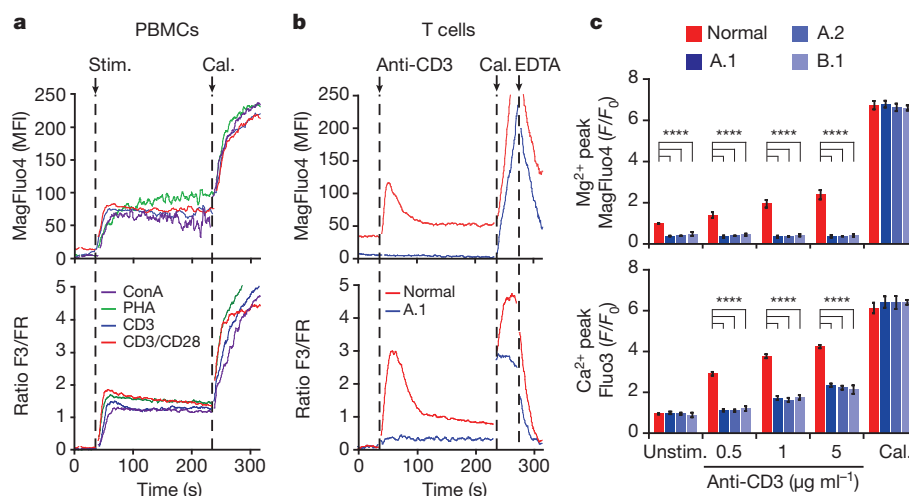
in  $\sim 90\%$  of target regions. This revealed a 10-base-pair (bp) deletion in the two brothers present in MAGT1 (Fig. 2c, Supplementary Fig. 3a), a gene encoding a magnesium transporter<sup>34–36</sup>. This deletion was not detected in the mother's complementary DNA or 100 normal individuals (data not shown). The deletion removes a splice donor site located in the 3' exon–intron junction of exon 7 and was present in the grandmother and great-grandmother of the patients (Fig. 2a, c, Supplementary Fig. 3a). The patients' mutant MAGT1 splice variant was  $\sim 150$  bp smaller than the mother's normal splice variant of approximately 1,100 bp and missing both exon 7 and 8, leading to a premature stop codon (Supplementary Fig. 3b, c). Apparent nonsense-mediated decay caused markedly decreased messenger RNA expression (Fig. 2d). The MAGT1 protein was undetectable in the patients' cells by western blot or immunofluorescent cell surface staining (Fig. 2e, f).

Additional screening yielded another immunodeficient patient (B.1) with chronic Epstein–Barr virus infection and a nonsense mutation in exon 3 of MAGT1, leading to a 90% decrease in mRNA expression (Fig. 2a, c, d, Supplementary Table 1, Supplementary Fig. 4a). The patient died five years ago from chronic Epstein–Barr virus-associated lymphoma at the age of 45. Like the other two patients, patient B.1 exhibited a similar T-cell defect in NF- $\kappa$ B and NFAT nuclear translocation in response to TCR but not PMA/ionomycin stimulation (Supplementary Fig. 4b, c).

### TCR-induced Mg<sup>2+</sup> and Ca<sup>2+</sup> influx defects

MAGT1 is a mammalian Mg<sup>2+</sup>-selective transporter whose physiological function is not well understood<sup>34,35</sup>. The full-length protein of 367 amino acids encodes a signal peptide, a large amino-terminal segment, four transmembrane domains, and a small carboxy-terminal tail with little similarity to any other known transporter except TUSC3, a non-selective Mg<sup>2+</sup> transporter<sup>36</sup>. We measured ion uptake in normal and patients' lymphocytes using fluorescent probes sensitive for Ca<sup>2+</sup> (Fluo3-AM and Fura red-AM) or Mg<sup>2+</sup> (Magfluor4-AM), which exhibited no detectable cross-reactivity (Supplementary Fig. 5a). Initial experiments showed a low basal level





**Figure 3 | TCR stimulation induces a MAGT1-dependent Mg<sup>2+</sup> influx.** **a–c**, Flux of Mg<sup>2+</sup> (MagFluo4, upper panels) and Ca<sup>2+</sup> (ratio Fluo3/Fura Red (F3/FR), lower panels); see Methods for details. MFI, mean fluorescence intensity. **a**, Fluxes in normal PBMCs stimulated (Stim.) with ConA, PHA, anti-CD3, or anti-CD3/anti-CD28. **b**, Fluxes in healthy control T cells or patient A.1 after 1  $\mu\text{g ml}^{-1}$  anti-CD3 stimulation. In **a** and **b**, maximum fluxes were

detected with calcimycin (Cal.) treatment to control for dye loading. In **b**, fluxes were quenched by EDTA (EDTA) treatment to show specificity. **c**, Peak value of the fluxes in healthy control T cells and the three patients on stimulation with indicated anti-CD3 concentrations.  $F/F_0$ , fluorescence normalized to the value at time 0. Error bars, s.e.m. ( $n = 3$ ); \*\*\*\* $P < 0.0001$ .

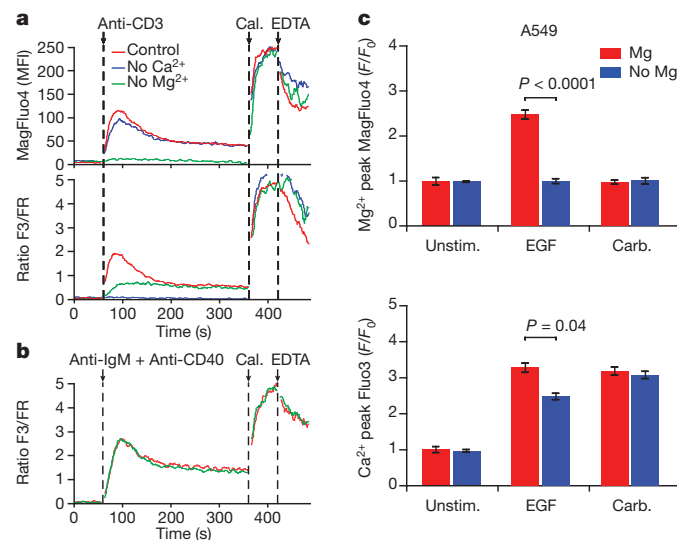
of free Mg<sup>2+</sup> and defective passive Mg<sup>2+</sup> influx, whereas that for Ca<sup>2+</sup> was normal (Supplementary Fig. 5a). The total Ca<sup>2+</sup> and Mg<sup>2+</sup> levels in the patients' T cells determined by inductively coupled plasma mass spectrometry were normal, indicating that MAGT1 deficiency chiefly affects free Mg<sup>2+</sup> and that general metabolic processes requiring bound Mg<sup>2+</sup> should not be affected (Supplementary Fig. 5b).

We next examined whether various TCR stimuli would affect free Mg<sup>2+</sup> transport. We observed a robust transient Mg<sup>2+</sup> influx together with the well-documented Ca<sup>2+</sup> influx in normal PBMCs stimulated with various TCR agonists (Fig. 3a). The apparent Mg<sup>2+</sup> influx was not due to cross-detection of the Ca<sup>2+</sup> influx, because the specific Ca<sup>2+</sup> chelator 1,2-bis(*o*-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid acetoxymethyl ester (BAPTA-AM) abolished the Ca<sup>2+</sup> fluorescence but not the Mg<sup>2+</sup> fluorescence (Supplementary Fig. 6a). The Mg<sup>2+</sup> influx was not detectable in the patients' T cells even with the strongest TCR agonist tested (5  $\mu\text{g ml}^{-1}$  anti-CD3) (Fig. 3b, c). Surprisingly, we also found that the Ca<sup>2+</sup> influx was severely compromised in the patients' T cells across a broad dose range of anti-CD3 stimulation (Fig. 3b, c). The TCR-induced Mg<sup>2+</sup> influx was selective, as stimulation of T cells with secondary lymphoid tissue chemokine (SLC/CXCL21), Fas-ligand (FasL) and tumour necrosis factor- $\alpha$  (TNF $\alpha$ ) caused no Mg<sup>2+</sup> influx and the Ca<sup>2+</sup> influxes induced by SLC and FasL were normal in the patient's cells (Supplementary Fig. 6b). Also, no Mg<sup>2+</sup> influx was discernible in B lymphocytes following anti-IgM and anti-CD40 stimulation, and the ensuing Ca<sup>2+</sup> influx was not diminished in patients' B cells (Supplementary Fig. 6c). Although the patients' B cells exhibited reduced basal free Mg<sup>2+</sup>, their B-cell activation was normal (Fig. 1d).

We next examined the hypothesis that the defective Ca<sup>2+</sup> influx in the patients was secondary to the loss of the Mg<sup>2+</sup> influx. We first explored the relationship between the TCR-stimulated influxes in normal T cells by modulating [Mg<sup>2+</sup>]<sub>e</sub> and [Ca<sup>2+</sup>]<sub>e</sub> in the extracellular buffer. Both Mg<sup>2+</sup> and Ca<sup>2+</sup> influxes were optimal when [Mg<sup>2+</sup>]<sub>e</sub> and [Ca<sup>2+</sup>]<sub>e</sub> = 1 mM, respectively, but were abolished when [Mg<sup>2+</sup>]<sub>e</sub> and [Ca<sup>2+</sup>]<sub>e</sub> = 0, respectively (Fig. 4a). Moreover, when [Mg<sup>2+</sup>]<sub>e</sub> = 0, the Ca<sup>2+</sup> influx was decreased, but when [Ca<sup>2+</sup>]<sub>e</sub> = 0, the Mg<sup>2+</sup> influx was unaffected (Fig. 4a). These results show that the TCR-induced Ca<sup>2+</sup> influx is partially dependent on [Mg<sup>2+</sup>]<sub>e</sub>. Thus, we infer that the Ca<sup>2+</sup> influx defect in T cells associated with MAGT1 deficiency is secondary to the loss of the TCR-stimulated Mg<sup>2+</sup> influx. Conversely, when [Mg<sup>2+</sup>]<sub>e</sub> = 0, the Ca<sup>2+</sup> influx in B cells after the BCR stimulation is normal, which is consistent with the absence of BCR-induced

Mg<sup>2+</sup> influxes (Fig. 4b, Supplementary Fig. 6b). Thus, the loss of the TCR-induced Ca<sup>2+</sup> influx in the absence of extracellular Mg<sup>2+</sup> is not likely to be due to a deficiency in Mg<sup>2+</sup> as a cofactor required for ATP-dependent processes, as B cells had no comparable defect. This is underscored by the fact that both B and T cells from the MAGT1-deficient patients had the same partial reduction in the free [Mg<sup>2+</sup>]<sub>i</sub>, but only the latter exhibited functional defects in antigen receptor signalling.

To determine whether Mg<sup>2+</sup>-promoted Ca<sup>2+</sup> influxes are important in other tissues, we treated two transformed epithelial cell lines, A549 and HepG2, with epidermal growth factor (EGF), which is known to cause a PLC $\gamma$ 1-dependent Ca<sup>2+</sup> influx analogous to that induced by TCR engagement<sup>37</sup>. In each cell type, we observed a clear



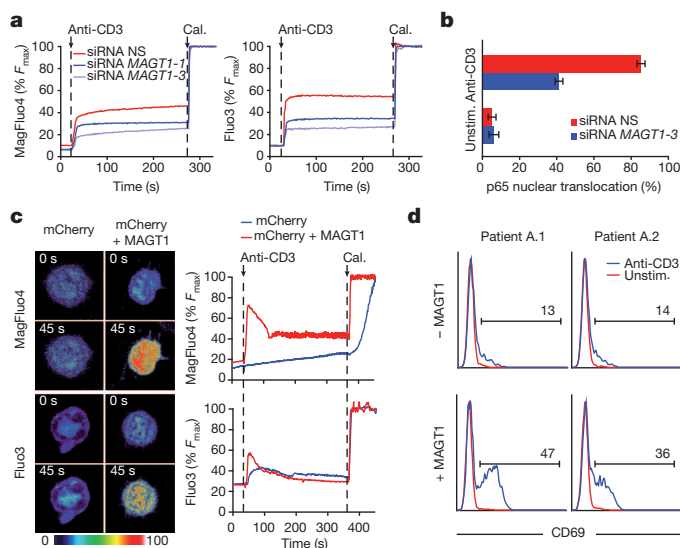
**Figure 4 | Requirement of receptor-stimulated Mg<sup>2+</sup> influx for Ca<sup>2+</sup> influx.** **a**, Mg<sup>2+</sup> (MagFluo4, upper panel) and Ca<sup>2+</sup> (ratio F3/FR, lower panel) flux in healthy control T cells stimulated with anti-CD3 in buffers containing 1 mM Mg<sup>2+</sup> and Ca<sup>2+</sup> (control) or lacking either ion. **b**, Ca<sup>2+</sup> flux in B cells stimulated with anti-IgM and anti-CD40. In **a** and **b**, calcimycin (Cal.) and EDTA addition display maximum influx and ion chelation, respectively. **c**, Graphs represent the fold change of the peak of Mg<sup>2+</sup> (upper panels) and Ca<sup>2+</sup> (lower panels) flux in A549 cells either unstimulated (Unstim.) or stimulated with epidermal growth factor (EGF) or carbachol (Carb.). Error bars, s.e.m. ( $n = 3$ ).

dose-dependent influx of  $Mg^{2+}$  induced by EGF that was abrogated when  $[Mg^{2+}]_e = 0$  (Fig. 4c, Supplementary Fig. 7). The EGF-induced  $Ca^{2+}$  influx was also decreased, though not abolished, by  $Mg^{2+}$  depletion. By contrast, carbachol, an agonist for acetylcholine receptors that induces a  $Ca^{2+}$  influx through PLC $\beta$ , failed to induce a  $Mg^{2+}$  influx. In addition, the carbachol-induced  $Ca^{2+}$  influx was not altered when  $[Mg^{2+}]_e = 0$ . Thus, receptor-induced  $Mg^{2+}$  influxes that promote  $Ca^{2+}$  influxes occur in both lymphoid and non-lymphoid tissue types. Moreover, receptors that trigger a  $Ca^{2+}$  influx through PLC $\gamma$ 1, but not PLC $\beta$  or PLC $\gamma$ 2 (BCR), can induce a  $Mg^{2+}$  influx to regulate the  $Ca^{2+}$  influx.

### Knockdown and reconstitution of MAGT1

To demonstrate that decreased MAGT1 expression can account for the immunological and signalling defects observed in the patients, we knocked down MAGT1 in normal human T cells by transient short interfering RNA (siRNA) transfections. We found that the TCR-stimulated  $Mg^{2+}$  and  $Ca^{2+}$  influxes were decreased in proportion with the degree of MAGT1 mRNA knockdown (Fig. 5a, Supplementary Fig. 8a). Similar to the patients' T-cell phenotype, MAGT1 knockdown impeded TCR-induced p65 nuclear translocation (Fig. 5b, Supplementary Fig. 8b). These results verify that MAGT1 is necessary for normal T-cell activation.

To determine whether MAGT1 deficiency is sufficient to explain the patients' functional defects, we reconstituted MAGT1 expression by lentiviral transduction of patient T cells. Positively transduced cells marked by a coexpressed fluorescent marker (mCherry) were examined by live cell confocal imaging of the influxes. We found that expressing wild-type MAGT1 in the patients' T cells restored a TCR-stimulated  $Mg^{2+}$  influx (Fig. 5c). MAGT1 restoration also improved the TCR-stimulated  $Ca^{2+}$  influx, thereby validating our conjecture that it is contingent on the  $Mg^{2+}$  influx (Fig. 5c). The expression of MAGT1 also augmented other activation events of the patients' T cells, such as TCR-induced CD69 upregulation (Fig. 5d). Thus, MAGT1 is necessary and sufficient for the  $Mg^{2+}$  influx required for optimal T-cell activation,



**Figure 5 | Knockdown and rescue of MAGT1.** Healthy T cells transfected with non-specific (NS) or MAGT1 siRNAs. **a**,  $Mg^{2+}$  (MagFluo4, left) and  $Ca^{2+}$  (Fluo3, right) flux upon anti-CD3 stimulation. **b**, Percentage of nuclear p65 after MAGT1 knockdown. Error bars represent s.e.m. ( $n = 3$ ). **c**, Time-lapse imaging (left; time shown in seconds) or cytometry (right) of  $Mg^{2+}$  (upper) and  $Ca^{2+}$  (lower) flux in T cells transduced with lentiviruses expressing mCherry or mCherry + MAGT1. **d**, Flow cytometry of CD69 expression on CD4<sup>+</sup> T cells transduced with lentiviruses expressing MAGT1 or not and either unstimulated (Unstim.) or stimulated with anti-CD3. Percentage of cells are shown for the indicated gates. Calcimycin (Cal.) addition displays maximum influx. %  $F_{max}$  fluorescence normalized as percentage of the maximum value.

and MAGT1 deficiency is the proximate cause of the T cell activation defect in this primary immunodeficiency.

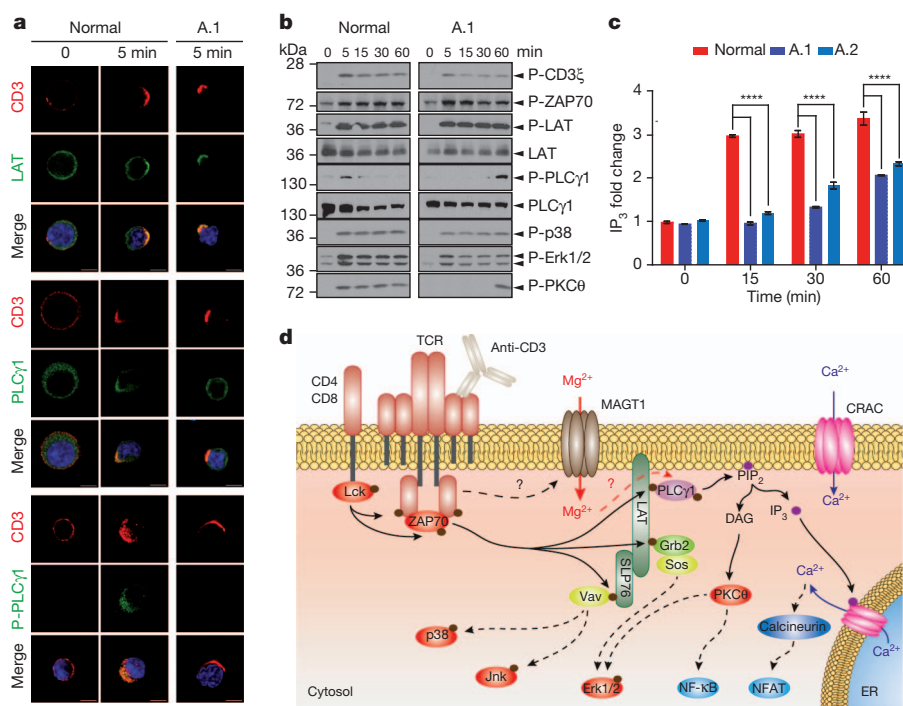
### Loss of MAGT1 impairs PLC $\gamma$ 1 activation

To understand the molecular mechanism of the  $Ca^{2+}$  influx defect observed in MAGT1-deficient patients, we examined proximal TCR signalling components (Fig. 6d). TCR engagement causes clustering and phosphorylation of the CD3 $\zeta$  chain by the Src-family leukocyte-specific protein tyrosine kinase (Lck) and subsequent recruitment of the protein tyrosine kinase ZAP70, which phosphorylates the scaffold proteins linker of activated T cells (LAT) and Src homology 2 (SH2) domain-containing leukocyte protein of 76 kDa (SLP76)<sup>20,38</sup>. These phosphorylated scaffolds then bind inducible T cell kinase (Itk), which activates PLC $\gamma$ 1 and thereby generates inositol triphosphate (IP<sub>3</sub>) and diacylglycerol (DAG) as second messengers to trigger  $Ca^{2+}$  mobilization and protein kinase C $\theta$  (PKC $\theta$ ) activation, respectively. The  $Ca^{2+}$  influx modulates the protein phosphatase calcineurin, which, together with PKC $\theta$ , activates downstream transcription factors such as NF- $\kappa$ B and NFAT<sup>39</sup>. We found that TCR cluster formation and LAT and PLC $\gamma$ 1 recruitment to these clusters were intact in patient T cells after anti-CD3 stimulation (Fig. 6a, Supplementary Fig. 9). Moreover, early TCR activation events including the phosphorylation of CD3 $\zeta$ , Lck, ZAP70 and LAT induced by TCR ligation were normal in the patient's T cells using flow cytometric staining (Supplementary Fig. 10) and western blots (Fig. 6b). Again, this normal series of activation events excludes a general defect in  $Mg^{2+}$  as a co-factor for ATP-requiring processes accounting for the TCR signalling defects in MagT1-deficient T cells. By contrast, PLC $\gamma$ 1 activation was markedly delayed by almost one hour in the patient T cells compared to healthy control T cells (Fig. 6b). Moreover, the activating phosphorylation of PKC $\theta$  and IP<sub>3</sub> generation downstream of PLC $\gamma$ 1 were significantly reduced (Fig. 6b, c). On the other hand, we found that TCR signalling events that do not require PLC $\gamma$ 1, such as the phosphorylation of the mitogen activated protein kinases (MAPKs) p38 and Erk1/2, were intact in the patient T cells (Fig. 6b). This deficiency in PLC $\gamma$ 1 and PKC $\theta$  activation following TCR stimulation was recapitulated by RNAi silencing of MAGT1 in normal T lymphocytes (Supplementary Fig. 11). Thus,  $Mg^{2+}$  can regulate signal transduction pathways involving PLC $\gamma$ 1 in lymphoid and non-lymphoid cells, and MAGT1-deficiency profoundly delays the activation of the PLC $\gamma$ 1 branch of TCR signalling.

### Discussion

We have found a biologically important transient  $Mg^{2+}$  influx mediated by MAGT1 during T-cell activation and EGF stimulation of epithelial cells. Three fundamental features of a second messenger have been put forward<sup>40</sup>: (1) its levels must increase rapidly in response to a stimulus, which is typically the engagement of a cell-surface receptor (first messenger), (2) it needs to alter the rate of one or more cellular processes, and (3) it exerts cell-type specific activity because different cells harbour different complements of enzymes. We find that  $Mg^{2+}$  fulfils these requirements and is a kinetic regulator of signalling in T cells. We found no difference in total  $Mg^{2+}$  content between our patients and healthy controls, indicating that MAGT1 does not influence general  $Mg^{2+}$  homeostasis or its cofactor function. Rather, we measured rapid and transient free  $Mg^{2+}$  influxes after TCR engagement which depends on the MAGT1 transporter. These transient  $Mg^{2+}$  influxes increased the rate of PLC $\gamma$ 1 activation and the corresponding  $Ca^{2+}$  influx in T cells. B cells, which depend on PLC $\gamma$ 2 for the BCR-induced  $Ca^{2+}$  influx, have no such requirement for MAGT1-mediated  $Mg^{2+}$  influxes despite their expression of MAGT1. Because ions diffuse rapidly as second messengers, a  $Mg^{2+}$  influx may promote rapid spatial integration of antigen and costimulatory receptor signals critical for T-cell activation. Our findings shed new light on older observations that extracellular magnesium promotes lymphocyte activation synergistically with calcium by T-cell, but not B-cell,





**Figure 6 | MAGT1 deficiency impairs PLC $\gamma$ 1 activation upon TCR stimulation.** **a**, Confocal images of TCR clustering induced by anti-CD3. Cells were stained for LAT, PLC $\gamma$ 1 or phospho (P)-PLC $\gamma$ 1. Scale bar, 5  $\mu$ m. **b**, **c**, Immunoblot of the indicated signalling proteins and phosphoproteins (**b**) and quantification of cellular IP $_3$  level (**c**) in healthy control and patient's T

cells stimulated with anti-CD3 for indicated times. Error bars, s.e.m. ( $n = 3$ ); \*\*\*\* $P < 0.0001$ . **d**, Hypothetical schematic depicting how the MAGT1-mediated Mg $^{2+}$  influx participates in TCR signalling. Solid arrows indicate direct effects; dashed arrows indicate indirect effects. ER, endoplasmic reticulum.

mitogens, and that mice fed Mg $^{2+}$ -deficient diets have lower calcineurin activity in their splenocytes<sup>11–13,41,42</sup>. It will also be of interest to determine if defective thymopoiesis caused by disruption of the TRPM7 Mg $^{2+}$  channel could be due to the loss of the signalling function of Mg $^{2+}$  influxes<sup>43</sup>.

The patients examined in this study have a novel X-linked primary immunodeficiency impairing thymic production of CD4 $^{+}$  T cells and circulating T-cell function which we now denominate XMEN (X-linked immunodeficiency with magnesium defect and Epstein–Barr virus infection and neoplasia) disease. The defective development and function of T cells in XMEN patients categorizes them as T cell immunodeficiency patients, although they have a mild phenotype comprising chiefly uncontrolled viral infections consistent with a T-cell-specific defect<sup>44</sup>. They also share features with X-linked lymphoproliferative disease (XLP) except for the absence of NKT cell deficits or haemophagocytic lymphohistiocytosis<sup>45</sup>. Like XMEN patients, mice with PLC $\gamma$ 1-deficient T cells have fewer CD4 $^{+}$  T cells and defective TCR activation, but they also have features that we did not observe in the patients, such as impaired ERK activation, lower numbers of CD8 $^{+}$  T cells and FoxP3 $^{+}$  regulatory T cells, and inflammatory/autoimmune symptoms<sup>46</sup>. The fact that XMEN patients only have a kinetic block in PLC $\gamma$ 1 activation, whereas the mice have a complete deficiency, could account for these differences.

Finally, the integrity of early TCR activation events in the patients' T cells explains how MAGT1 acts as a TCR-gated transporter that controls a new step in the later contingent series of TCR signalling events<sup>47</sup> (Fig. 6d). The selective requirement of this Mg $^{2+}$  influx for the activation of T cells but not B cells suggests that MAGT1 may be a useful therapeutic target for diseases requiring T-cell-specific immunomodulation.

## METHODS SUMMARY

All human subjects in this study provided informed consent in accordance with Helsinki principles to be enrolled in research protocols approved by the institutional review board of the National Institute of Allergy and Infectious Diseases,

NIH. Patients' and healthy control PBMCs were Ficoll-purified and activated with anti-CD3 and anti-CD28 for 3 days and then continuously cultured in 10% RPMI media supplemented with 100 U ml $^{-1}$  IL-2 for 3–4 weeks. For assessment of primary stimulation, cells were collected and stained with anti-CD2, anti-CD4 and anti-CD8 antibodies, and various activation markers for 30 min at 4  $^{\circ}$ C and analysed with a FACS Calibur or a LSRII flow cytometer (BD Biosciences). Activated T cells were restimulated with various T-cell activation agents, and processed for immunofluorescence imaging, immunoblotting and influx assays, as described in Methods. Genomic DNA isolated from activated T cells was prepared for lysis, SureSelect Human X Chromosome (Agilent) target enrichment, single-read Solexa sequencing on Illumina Genome Analyser IIx, and Sanger sequencing according to the manufacturer's instructions. All  $P$  values were calculated with the Student's  $t$ -test using PRISM software (GraphPad Software), with a two-tailed distribution.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 30 March; accepted 3 June 2011.

1. Cakmak, I. & Kirkby, E. A. Role of magnesium in carbon partitioning and alleviating photooxidative damage. *Physiol. Planta* **133**, 692–704 (2008).
2. Cowan, J. A. Structural and catalytic chemistry of magnesium-dependent enzymes. *Biometals* **15**, 225–235 (2002).
3. Yang, W., Lee, J. Y. & Nowotny, M. Making and breaking nucleic acids: two-Mg $^{2+}$ -ion catalysis and substrate specificity. *Mol. Cell* **22**, 5–13 (2006).
4. Gasser, A., Bruhn, S. & Guse, A. H. Second messenger function of nicotinic acid adenine dinucleotide phosphate revealed by an improved enzymatic cycling assay. *J. Biol. Chem.* **281**, 16906–16913 (2006).
5. Grubbs, R. D. & Maguire, M. E. Magnesium as a regulatory cation: criteria and evaluation. *Magnesium* **6**, 113–127 (1987).
6. Murphy, E. Mysteries of magnesium homeostasis. *Circ. Res.* **86**, 245–248 (2000).
7. Permyakov, E. A. & Kretsinger, R. H. Cell signaling, beyond cytosolic calcium in eukaryotes. *J. Inorg. Biochem.* **103**, 77–86 (2009).
8. Takaya, J., Higashino, H. & Kobayashi, Y. Can magnesium act as a second messenger? Current data on translocation induced by various biologically active substances. *Magnesium* **13**, 139–146 (2000).
9. Hogan, P. G., Lewis, R. S. & Rao, A. Molecular basis of calcium signaling in lymphocytes: STIM and ORAI. *Annu. Rev. Immunol.* **28**, 491–533 (2010).
10. Romani, A. M. Magnesium homeostasis in mammalian cells. *Front. Biosci.* **12**, 308–331 (2007).

11. Abboud, C. N., Scully, S. P., Lichtman, A. H., Brennan, J. K. & Segel, G. B. The requirements for ionized calcium and magnesium in lymphocyte proliferation. *J. Cell. Physiol.* **122**, 64–72 (1985).
12. Modiano, J. F., Kelepouris, E., Kern, J. A. & Nowell, P. C. Requirement for extracellular calcium or magnesium in mitogen-induced activation of human peripheral blood lymphocytes. *J. Cell. Physiol.* **135**, 451–458 (1988).
13. Whitney, R. B. & Sutherland, R. M. The influence of calcium, magnesium and cyclic adenosine 3',5'-monophosphate on the mixed lymphocyte reaction. *J. Immunol.* **108**, 1179–1183 (1972).
14. Ng, L. L., Davies, J. E. & Garrido, M. C. Intracellular free magnesium in human lymphocytes and the response to lectins. *Clin. Sci. (Lond.)* **80**, 539–547 (1991).
15. Rijkers, G. T. & Griffioen, A. W. Changes in free cytoplasmic magnesium following activation of human lymphocytes. *Biochem. J.* **289**, 373–377 (1993).
16. Chun, H. J. *et al.* Pleiotropic defects in lymphocyte activation caused by caspase-8 mutations lead to human immunodeficiency. *Nature* **419**, 395–399 (2002).
17. Notarangelo, L. D. Primary immunodeficiencies. *J. Allergy Clin. Immunol.* **125**, S182–S194 (2010).
18. Zhang, Q. *et al.* Combined immunodeficiency associated with DOCK8 mutations. *N. Engl. J. Med.* **361**, 2046–2055 (2009).
19. Chan, A. C. *et al.* ZAP-70 deficiency in an autosomal recessive form of severe combined immunodeficiency. *Science* **264**, 1599–1601 (1994).
20. Peterson, E. J. & Koretzky, G. A. Signal transduction in T lymphocytes. *Clin. Exp. Rheumatol.* **17**, 107–114 (1999).
21. Arpaia, E., Shahar, M., Dadi, H., Cohen, A. & Roifman, C. M. Defective T cell receptor signaling and CD8<sup>+</sup> thymic selection in humans lacking zap-70 kinase. *Cell* **76**, 947–958 (1994).
22. Feske, S. *et al.* A mutation in Orai1 causes immune deficiency by abrogating CRAC channel function. *Nature* **441**, 179–185 (2006).
23. Unexplained CD4<sup>+</sup> T-lymphocyte depletion in persons without evident HIV infection — United States. *MMWR Morb. Mortal. Wkly. Rep.* **41**, 541–545 (1992).
24. Laurence, J., Siegal, F. P., Schattner, E., Gelman, I. H. & Morse, S. Acquired immunodeficiency without evidence of infection with human immunodeficiency virus types 1 and 2. *Lancet* **340**, 273–274 (1992).
25. Smith, D. K., Neal, J. J. & Holmberg, S. D. Unexplained opportunistic infections and CD4<sup>+</sup> T-lymphocytopenia without HIV infection. An investigation of cases in the United States. *N. Engl. J. Med.* **328**, 373–379 (1993).
26. Fauci, A. S. CD4<sup>+</sup> T-lymphocytopenia without HIV infection — no lights, no camera, just facts. *N. Engl. J. Med.* **328**, 429–431 (1993).
27. Freier, S. *et al.* Hereditary CD4<sup>+</sup> T lymphocytopenia. *Arch. Dis. Child.* **78**, 371–372 (1998).
28. Lin, S. J., Chao, H. C., Yan, D. C. & Kuo, M. L. Idiopathic CD4<sup>+</sup> T lymphocytopenia in two siblings. *Pediatr. Hematol. Oncol.* **18**, 153–156 (2001).
29. Lobato, M. N., Spira, T. J. & Rogers, M. F. CD4<sup>+</sup> T lymphocytopenia in children: lack of evidence for a new acquired immunodeficiency syndrome agent. *Pediatr. Infect. Dis. J.* **14**, 527–535 (1995).
30. Junge, S. *et al.* Correlation between recent thymic emigrants and CD31<sup>+</sup> (PECAM-1) CD4<sup>+</sup> T cells in normal individuals during aging and in lymphopenic children. *Eur. J. Immunol.* **37**, 3270–3280 (2007).
31. Kimmig, S. *et al.* Two subsets of naive T helper cells with distinct T cell receptor excision circle content in human adult peripheral blood. *J. Exp. Med.* **195**, 789–794 (2002).
32. Kohler, S. & Thiel, A. Life after the thymus: CD31<sup>+</sup> and CD31<sup>−</sup> human naive CD4<sup>+</sup> T-cell subsets. *Blood* **113**, 769–774 (2009).
33. Wengler, G. S. *et al.* A PCR-based non-radioactive X-chromosome inactivation assay for genetic counseling in X-linked primary immunodeficiencies. *Life Sci.* **61**, 1405–1411 (1997).
34. Goytain, A. & Quamme, G. A. Identification and characterization of a novel mammalian Mg<sup>2+</sup> transporter with channel-like properties. *BMC Genomics* **6**, 48–66 (2005).
35. Quamme, G. A. Molecular identification of ancient and modern mammalian magnesium transporters. *Am. J. Physiol. Cell Physiol.* **298**, C407–C429 (2010).
36. Zhou, H. & Clapham, D. E. Mammalian MagT1 and TUSC3 are required for cellular magnesium uptake and vertebrate embryonic development. *Proc. Natl Acad. Sci. USA* **106**, 15750–15755 (2009).
37. Xie, Z., Peng, J., Pennypacker, S. D. & Chen, Y. Critical role for the catalytic activity of phospholipase C-γ1 in epidermal growth factor-induced cell migration. *Biochem. Biophys. Res. Commun.* **399**, 425–428 (2010).
38. Weiss, A. & Littman, D. R. Signal transduction by lymphocyte antigen receptors. *Cell* **76**, 263–274 (1994).
39. Nel, A. E. T-cell activation through the antigen receptor. Part 1: signaling components, signaling pathways, and signal integration at the T-cell antigen receptor synapse. *J. Allergy Clin. Immunol.* **109**, 758–770 (2002).
40. Sutherland, E. W. Studies on the mechanism of hormone action. *Science* **177**, 401–408 (1972).
41. Flynn, A. Control of *in vitro* lymphocyte proliferation by copper, magnesium and zinc deficiency. *J. Nutr.* **114**, 2034–2042 (1984).
42. Sabbagh, F., Lecerf, F., Hulin, A., Bac, P. & German-Fattal, M. Effect of hypomagnesemia on allogeneic activation in mice. *Transpl. Immunol.* **20**, 83–87 (2008).
43. Jin, J. *et al.* Deletion of Trpm7 disrupts embryonic development and thymopoiesis without altering Mg<sup>2+</sup> homeostasis. *Science* **322**, 756–760 (2008).
44. Cossu, F. Genetics of SCID. *Ital. J. Pediatr.* **36**, 36–76 (2010).
45. Filipovich, A. H., Zhang, K., Snow, A. L. & Marsh, R. A. X-linked lymphoproliferative syndromes: brothers or distant cousins? *Blood* **116**, 3398–3408 (2010).
46. Fu, G. *et al.* Phospholipase Cγ1 is essential for T cell development, activation, and tolerance. *J. Exp. Med.* **207**, 309–318 (2010).
47. Crabtree, G. R. Contingent genetic regulatory events in T lymphocyte activation. *Science* **243**, 355–361 (1989).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank L. Zheng, A. Weiss, R. Germain, R. Siegel, F. Wolf and P. Schwartzberg for critically reading the manuscript; F. Wolf for advice on magnesium assessments; L. Zheng, C. Lowell and A. Weiss for advice on experiments and data; H. Jing for making HVS lines from patient cells; P. Chen for assistance with plasmid DNA preparation; N. Sandler for flow cytometry assistance; A. Snow and H. Jing for assistance with genomic DNA library preparation for Solexa sequencing; J. Almenara and Illumina staff for Solexa assistance; D. Killilea for assistance with MS-ICP data interpretation; and A. Irani for referring the patients. F.-Y.L. is in the Medical Scientist Training Program at the University of California–San Francisco and thanks K. Shannon and J. Toutolmin for support and encouragement. This work was supported by the Division of Intramural Research of the National Institute of Allergy and Infectious Diseases of the US National Institutes of Health.

**Author Contributions** F.-Y.L. characterized the MAGT1 mutations and the TCR activation defect in the patients. B.C.-D. characterized the Mg<sup>2+</sup> influx and the signalling defects. B.C.-D., F.-Y.L., H.C.S. and M.J.L. conceived and planned the experiments, and prepared the manuscript. J.C.D. performed the lyonization assay. C.K. performed the RT-PCR experiments. G.U., J.I.C. and H.C.S. referred patients and provided clinical data. H.F.M. coordinated clinical protocol and sample collection. D.C.D. provided assistance with sequencing and flow cytometry, and guided some patient assessments. All authors discussed and revised the manuscript.

**Author Information** The Illumina sequencing data has been deposited in dbGaP with accession code phs000365.v1.p1. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.J.L. ([ilenardo@nih.gov](mailto:ilenardo@nih.gov)).



## METHODS

**Cells and reagents.** PBMCs were isolated from whole blood by Ficoll-Hypaque (GE Healthcare) density gradient centrifugation, washed, and resuspended at  $10^6$  cells  $\text{ml}^{-1}$  in complete RPMI 1640 medium (Lonza) containing 10% fetal calf serum (FCS), 2 mM glutamine, and penicillin and streptomycin (100 U  $\text{ml}^{-1}$  each, Invitrogen). T cells were activated with  $1 \mu\text{g ml}^{-1}$  each of anti-CD3 $\epsilon$  and anti-CD28 monoclonal antibodies (BD Biosciences) or with 0.8 ng  $\text{ml}^{-1}$  PMA (Sigma) and 0.4  $\mu\text{M}$  ionomycin (Sigma). After 3 days, activated T cells were washed and then cultured in complete RPMI-1640 medium supplemented with 100 U  $\text{ml}^{-1}$  rhIL-2 (R&D). B cells were enriched by negative selection using a B cell isolation kit II (Miltenyi).

Antibodies against p65 (RelA), phospho-PKC $\theta$  (S695) and phospho-CD3 $\zeta$  (C415.9A) were from Santa Cruz. Antibodies against CD4, CD8, CD31, CD40, CD69, CD86, hIgM and phospho-Lck (Y394) were from BD Bioscience. Antibodies against phospho-LAT (Y191), LAT, phospho-PLC $\gamma$ 1 (Y783), PLC $\gamma$ 1, phospho-p38 and phospho-Erk1/2 (T202/204) were from Cell Signaling Technology. ConA and PHA were from Sigma. Recombinant FasL and TNF $\alpha$  were from Alexis and recombinant SLC was from PeproTech. All flow cytometry staining and analysis were performed using standard protocols<sup>18,48</sup>.

**Lyonization assay.** Lyonization assays were adapted from ref. 49. Briefly, genomic DNA (1  $\mu\text{g}$ ) samples were first digested with RsaI (NEB) and then split into two for treatment with or without HpaII (NEB). Each digestion was carried out for at least 2 h and heat inactivated at 65 °C for 20 min. Digested DNA was purified by QIAquick PCR purification kit (Qiagen) and PCR amplified using 6-FAM-conjugated primers for the HUMARA locus: 5'-TGCGCGAAGTGATCAGAACC-3', 5'-TGGGCTTGGGGAGAACCATCC-3'. The amplified products were analysed on a 310 Gene Analyser (ABI) as recommended using ROX500 (ABI) as size standard.

**Exome sequencing and analysis.** Genomic DNA isolated from T cells by DNeasy Blood and Tissue kit (Qiagen) was processed using Genomic DNA Sample Prep Kit (Illumina) for SureSelect Human X Chromosome (Agilent) target enrichment according to manufacturer instructions. Captured DNA was subject to cluster generation using Single-Read Cluster Generation Kit v2 (Illumina) and massive parallel sequencing using SBS Sequencing Kit v3 (Illumina) on the Illumina Genome Analyser IIX. The data were imported into CLCbio Genomics Workbench software and aligned to the human genome reference (hg19) annotated with SNPs (130) downloaded from UCSC Genomics browser. Detection of SNPs and deletion insertion polymorphisms (DIPs) was performed on alignments in the software, and genes with missing coverage were identified using the targeted resequencing plug-in tool (CLCbio). Candidate genes were filtered using available filters on the software and in Excel spreadsheets. Sanger sequencing was performed as described previously<sup>18</sup>.

From the Illumina sequencing data, all single nucleotide polymorphisms (SNPs) and 1–2 bp DIPs that were not present in both brothers, not heterozygous in the mother, or that were found in her cDNA were excluded. Large DIPs (>1 kb) were screened by using comparative genomic hybridization 244K arrays (Agilent) as described previously<sup>18</sup>. To detect 3–1,000 bp DIPs, the targeted resequencing plug-in tool (CLCbio) was used to identify genes with missing coverage in the patients but not in the mother. SNPs and DIPs consistent with these search criteria were ruled out as common polymorphisms in the normal population by Sanger sequencing 100 normal controls.

**Transfection and transduction.** Full-length *MAGT1* was PCR amplified from normal control T-cell cDNA and subcloned into the BamI and NotI restriction sites of LeGO-iG<sup>50</sup> plasmid and into XhoI and BamHI restriction sites of the pLenti-bicistronic (Applied Biological Materials, ABM) plasmid. mCherry was also subcloned into the SnaBI and NotI restriction sites of pLenti-bicistronic plasmid. These constructs were each cotransfected with lentiviral second generation packaging mix (ABM) in 1:1 mix into HEK293T (70% confluent) using Turbofect transfection reagent (Fermentas). After 2 days, virus supernatants were harvested and concentrated by overnight ultracentrifugation at 20,000g at 4 °C. Virus pellets were resuspended in 100 RPMI and stored at –80 °C. Unstimulated PBMCs purified from whole blood by Ficoll were transduced using virus made from LeGO-iG as described<sup>10</sup>. Three days after transduction, half of the transduced PBMCs were activated with 1  $\mu\text{g ml}^{-1}$  OKT3 and analysed for activation by flow cytometry. Day 4 stimulated T cells were transduced similarly using virus made from pLenti-bicistronic vector. Two days after transduction, cells were harvested for influx measurements by confocal imaging.

For RNAi gene silencing of *MAGT1*, total T cells purified by Pan T cell isolation kit (Miltenyi Biotech) were transfected using P3 Primary Cell 4D-Nucleofector X kit (Lonza) with either Stealth RNAi Negative Control Duplexes or Stealth siRNA targeting *MAGT1* (Invitrogen): siMAGT1-1 (5'-GCCCAAAGAAAGAGGAGAUGGUGU-3'), siMAGT1-3 (5'-UCAUGUUCACUGCUCUCCAACUGCA-3'). Transfected cells were cultured in 10% RPMI media for at least 4 h before stimulation with 1  $\mu\text{g ml}^{-1}$  each of anti-CD3 and anti-CD28 for 48 h. Transfection was repeated again with the same siRNA, and transfected cells were cultured in 10%

RPMI media supplemented with 100 U  $\text{ml}^{-1}$  rhIL-2 for 48–72 h. Subsequently, cells were harvested for fluorimetric influx measurements and microscopic assessment of p65 translocation following stimulation with OKT3 or PMA/ionomycin. At the same time, cells were processed for real-time assessment of *MAGT1* mRNA expression using Taqman Gene Expression Assays (ABI).

**Confocal imaging.** For nuclear translocation of p65 and NFAT, cells were stimulated with either OKT3 or PMA/ionomycin for 30 min. Then cells were dropped on poly-L-lysine coated slides and fixed with 3% paraformaldehyde in phosphate-buffered saline (PBS), permeabilized with 0.05% Triton X-100 for 3 min at room temperature, and blocked with PBS containing 10% FCS. For TCR clustering, activated peripheral blood lymphocytes (PBL) from normal control and patient were incubated on ice with anti-CD3 antibody (1  $\mu\text{g ml}^{-1}$ ) for 30 min, washed and stained with Alexa-568 goat anti-mouse antibody for 30 min. Cell were then incubated at 37 °C for indicated times on poly-L-lysine coated slides and then fixed and permeabilized. All slides were stained with indicated primary antibody in 0.5% BSA-PBS for 45 min at room temperature. Slides were washed and incubated with an Alexa 488-conjugated donkey anti-rabbit antibodies (Invitrogen) for 45 min. Nuclei were stained with Hoechst 33342 (50 ng  $\text{ml}^{-1}$ , Invitrogen). Slides were washed in PBS, rinsed, and mounted with a coverslip using Fluoromount-G (Southern Biotechnology). All images were collected on a Leica TCS-NT/SP5 confocal microscope (Leica Microsystems) using a 63 $\times$  oil immersion objective NA 1.32, 'zoom X'. For enumeration of percentage of cells with nuclear NF- $\kappa$ B, NFAT or CD3 clustering, 450–700 cells were scored visually by a single observer.

**Calcium and magnesium flux experiments.** Cells were loaded with 1  $\mu\text{M}$  Magfluor4-AM (Invitrogen) or 1  $\mu\text{M}$  Fluo3-AM (Invitrogen) and 1  $\mu\text{M}$  Fura Red-AM (Invitrogen) for 20 min at 37 °C. Then, cells were washed in incubation buffer (1B: 120 mM N-methyl-D-glucamine, 20 mM HEPES, 4.7 mM KCl, 1.2 mM  $\text{KH}_2\text{PO}_4$ , 10 mM glucose, pH 7.4) with or without 1.2 mM  $\text{CaCl}_2$  and/or 1.2 mM  $\text{MgSO}_4$ . The loading efficiency was assessed by adding 1  $\mu\text{M}$  calcimycin, a  $\text{Ca}^{2+}$ /Mg $^{2+}$  ionophore that causes large influxes of both ions. For flow cytometry experiments, intracellular calcium and magnesium were measured using a BD LSRII flow cytometer. Kinetic analyses used the FlowJo software package (TreeStar), with percentage of responding cells defined as >95th percentile of unstimulated baseline. For fluorimeter experiments, cells were seeded in a 96-well plate and assessed with a POLARstar OPTIMA plate reader (BMG LabTech). For confocal microscopy, cells were plated on LabTek II Chamber Slide (Nunc) coated with 0.01% poly-L-lysine, and analysed with a Leica TCS-NT/SP5 confocal microscope (Leica Microsystems) using a 63 $\times$  oil immersion objective NA 1.32, 'zoom X'.

**Immunoblotting.** Activated T cells were starved for 4 h in RPMI medium without FCS. Then cells were stimulated with OKT3 (1  $\mu\text{g ml}^{-1}$ ) for indicated times. Stimuli were terminated by addition of ice-cold PBS, and pelleted cells were immediately lysed in 1% Triton X-100, 1% NP40, 50 mM Tris pH 8, 150 mM NaCl, 20 mM EDTA, 1 mM  $\text{Na}_2\text{VO}_4$ , 1 mM NaF, phosphatase inhibitor cocktail (Sigma) and complete protease inhibitor cocktail (Roche). Protein concentration was quantitated by BCA assay (Pierce). 30  $\mu\text{g}$  of cell lysates were separated by SDS-PAGE and transferred on PVDF membrane (Millipore). Membrane was blocked with BSA for 1 h before incubating with primary antibodies overnight. HRP-conjugated secondary antibodies were from Southern Biotechnology. Pierce Supersignal West chemiluminescent substrates were used for detection. *MAGT1* blotting was performed on unstarved activated T cells using 1:1,000 anti-MAGT1 (Proteintech).

**Quantitative total elemental content measurement.** 10 million CD8 $^{+}$  T cells purified from PBMCs using human CD8 Microbeads (Miltenyi Biotech) were stimulated with anti-CD3 and anti-CD28 for 3 days and subsequently cultured continuously in 10% RPMI media with 100 U  $\text{ml}^{-1}$  IL-2 at ~1–2 million cells per ml. When sufficient cells accumulated, duplicate aliquots of 5, 10 and 15 million cells were pelleted and dried overnight at 32 °C. Total magnesium, calcium and sulphur content was measured for each cell pellet by inductively coupled plasma mass spectrometry as described<sup>51</sup>. Similar results for Mg $^{2+}$  were obtained for CD4 $^{+}$  T cells (data not shown).

**IP $_3$  measurement.** IP $_3$  levels from stimulated T cells were assessed using an IP $_3$  ELISA kit (Cusabio) according manufacturer's instructions.

**Statistical analysis.** P values were calculated with the Students *t*-test using PRISM software (GraphPad Software), with a two-tailed distribution.

48. Petrovas, C. *et al.* PD-1 is a regulator of virus-specific CD8 $^{+}$  T cell survival in HIV infection. *J. Exp. Med.* **203**, 2281–2292 (2006).
49. Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M. & Belmont, J. W. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am. J. Hum. Genet.* **51**, 1229–1239 (1992).
50. Weber, K., Bartsch, U., Stocking, C. & Fehse, B. A multicolor panel of novel lentiviral "Gene Ontology" (LeGO) vectors for functional gene analysis. *Mol. Ther.* **16**, 698–706 (2008).
51. Killilea, D. W. & Ames, B. N. Magnesium deficiency accelerates cellular senescence in cultured human fibroblasts. *Proc. Natl Acad. Sci. USA* **105**, 5768–5773 (2008).

# Alfvénic waves with sufficient energy to power the quiet solar corona and fast solar wind

Scott W. McIntosh<sup>1</sup>, Bart De Pontieu<sup>2</sup>, Mats Carlsson<sup>3</sup>, Viggo Hansteen<sup>2,3</sup>, Paul Boerner<sup>2</sup> & Marcel Goossens<sup>4</sup>

**Energy is required to heat the outer solar atmosphere to millions of degrees (refs 1, 2) and to accelerate the solar wind to hundreds of kilometres per second (refs 2–6). Alfvén waves (travelling oscillations of ions and magnetic field) have been invoked as a possible mechanism to transport magneto-convective energy upwards along the Sun’s magnetic field lines into the corona. Previous observations<sup>7</sup> of Alfvénic waves in the corona revealed amplitudes far too small ( $0.5 \text{ km s}^{-1}$ ) to supply the energy flux ( $100\text{--}200 \text{ W m}^{-2}$ ) required to drive the fast solar wind<sup>8</sup> or balance the radiative losses of the quiet corona<sup>9</sup>. Here we report observations of the transition region (between the chromosphere and the corona) and of the corona that reveal how Alfvénic motions permeate the dynamic and finely structured outer solar atmosphere. The ubiquitous outward-propagating Alfvénic motions observed have amplitudes of the order of  $20 \text{ km s}^{-1}$  and periods of the order of  $100\text{--}500 \text{ s}$  throughout the quiescent atmosphere (compatible with recent investigations<sup>7,10</sup>), and are energetic enough to accelerate the fast solar wind and heat the quiet corona.**

In the chromospheric region, just above the solar photospheric surface but below the corona, observations from the Hinode spacecraft<sup>11</sup> revealed the presence of Alfvénic (Supplementary Information) waves that have significantly higher amplitudes ( $20 \text{ km s}^{-1}$ ) than in the corona ( $0.5 \text{ km s}^{-1}$ ). The chromospheric waves were observed on spicules<sup>10</sup>, which are jet-like features at the solar limb that protrude into the hot corona<sup>12,13</sup>. The apparent discrepancy between the chromospheric and coronal measurements<sup>7,10</sup> has raised concerns that these low-frequency Alfvénic motions do not contribute significantly to the energy balance of the outer solar atmosphere and inner heliosphere, perhaps because the large chromospheric wave energy flux is dissipated or reflected before reaching the corona.

Previous reports of Alfvén waves in the coronal plasmas<sup>7</sup>, which are at temperatures of  $\sim 10^6 \text{ K}$ , were based on measurements of line-of-sight (Doppler) velocity or, more indirectly, on non-thermal line broadening of spectral lines at lower spatial ( $\sim 4.5 \text{ Mm}$ ) and temporal ( $30 \text{ s}$ ) resolution. Here we use the He II 304-Å and Fe IX 171-Å channels of the Atmospheric Imaging Assembly (AIA<sup>14</sup>) on board the Solar Dynamics Observatory (SDO) satellite to make observations of the solar transition region and of the corona. The spatial resolution ( $\sim 870 \text{ km}$  on the Sun) and temporal resolution ( $8 \text{ s}$ ) available allow us to directly image the transverse swaying of magnetic field lines as Alfvénic waves pass through plasma at  $\sim 10^5 \text{ K}$  (in the transition region) or at coronal temperatures.

Movies of the 304-Å channel at the solar limb show a transition region that is dominated by spicular jets that shoot rapidly ( $20\text{--}150 \text{ km s}^{-1}$ ) upwards, often reaching heights of  $20,000 \text{ km}$  above the solar limb (Supplementary Movie 1). Movies of the same region in the 171-Å channel reveal coronal disturbances that propagate outward at high speeds ( $100\text{--}200 \text{ km s}^{-1}$ ; Supplementary Movie 2). Recent analysis<sup>15</sup> has showed that these spicules and propagating disturbances are the transition-region and coronal counterparts of the chromospheric spicules<sup>10,13,16</sup>, illustrating

that those spicules are associated with material that is heated to coronal temperatures. These transition-region and coronal features undergo significant Alfvénic motion, with displacements varying sinusoidally in time (Fig. 1, Supplementary Movies 1–4). Because the magnetic field dictates the direction of plasma motions at these heights of the solar atmosphere, the observed transverse (to the long axis of the ejected material) motions imply the presence of Alfvénic waves.

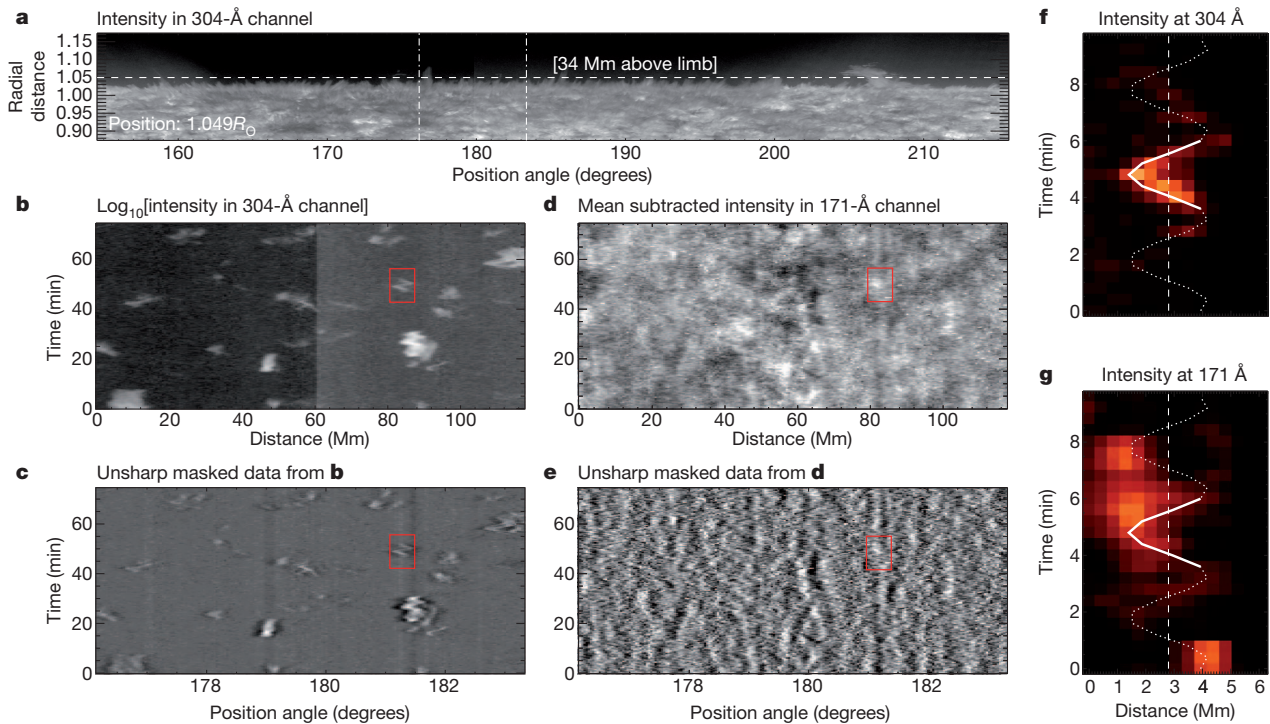
The SDO/AIA image sequences show a hot outer atmosphere that is replete with Alfvénic waves (Fig. 2; Supplementary Movies 1–7). The waves are traced by structures that do not have particularly long lifetimes (of the order of  $50\text{--}500 \text{ s}$ ) compared to the wave periods, and are difficult to detect because of the enormous line-of-sight superposition in the atmosphere above the solar limb. To avoid these issues, we use Monte Carlo simulations (Supplementary Information) to study the statistical properties of the waves. Below  $20 \text{ Mm}$  in height, we see a predominance of linear motion and partial swings (as opposed to full swings) in the space–time plots, because of the short transition-region spicule lifetimes. In the hotter corona, extending further from the surface, the spicule-related propagating coronal disturbances<sup>15</sup> display a mix of superposition and more complete sinusoidal motions with the same period. This suggests that the motions visible in the transition region and corona share a common origin.

Our Monte Carlo simulations (Fig. 2) show that the observations are compatible with the presence of Alfvénic waves with amplitudes of  $25 \text{ km s}^{-1}$  in the coronal hole and  $20 \text{ km s}^{-1}$  in the quiet Sun. Visual comparison of the observations and simulations roughly limit the periods to a range between  $150$  and  $550 \text{ s}$ : waves with very short periods of the order of  $50 \text{ s}$  or with very low amplitudes ( $< 15 \text{ km s}^{-1}$ ) do not fit the data very well (Supplementary Figs 2 and 3). The active solar corona can be analysed using a similar combination of techniques. Figure 3 (and Supplementary Movies 5–7) illustrates that the entire active region complex is riddled with transverse motion. The active region waves have shorter periods ( $100\text{--}400 \text{ s}$ ) and considerably lower amplitudes ( $5 \pm 5 \text{ km s}^{-1}$ ). By cross-correlating space–time plots at increasing distance above the limb, or along the coronal loops, we can estimate the phase speed (and associated error; see Supplementary Information) of the observed Alfvénic motions (Fig. 4; Supplementary Movies 3 and 4).

Using the wave amplitude inferred from our Monte Carlo simulations ( $v = 20\text{--}25 \text{ km s}^{-1}$  at a height of  $15 \text{ Mm}$ ) and the phase speeds inferred from cross-correlation of the coronal emission ( $V_A = 200\text{--}250 \text{ km s}^{-1}$  at the same height), we conservatively estimate the wave energy flux ( $E_A$ ) present in the coronal hole and quiet Sun regions using the expression  $E_A = f\rho v^2 V_A$ , where  $f$  is the filling factor of the waves, and  $\rho$  is the mass density in the each of the solar domains. The cross-field coherence in the oscillating features we observe (Supplementary Figs 4 and 5), and numerical simulations showing pervasive waves throughout the solar atmosphere<sup>10</sup>, indicate that these waves are volume-filling ( $f \approx 1$ ; see Supplementary Information). Using order-of-magnitude estimates for the density of wave-carrying features in the

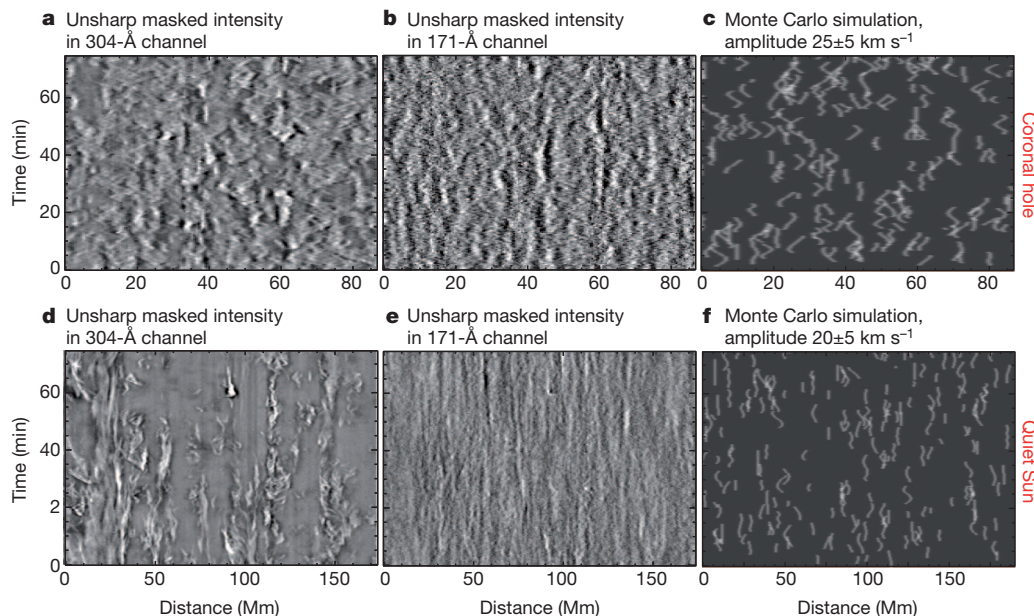
<sup>1</sup>High Altitude Observatory, National Center for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307, USA. <sup>2</sup>Lockheed Martin Solar and Astrophysics Laboratory, 3251 Hanover Street, Palo Alto, California 94304, USA. <sup>3</sup>Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, 0315 Oslo, Norway. <sup>4</sup>Department of Mathematics, Centre for Plasma Astrophysics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, B-3001 Heverlee, Belgium.





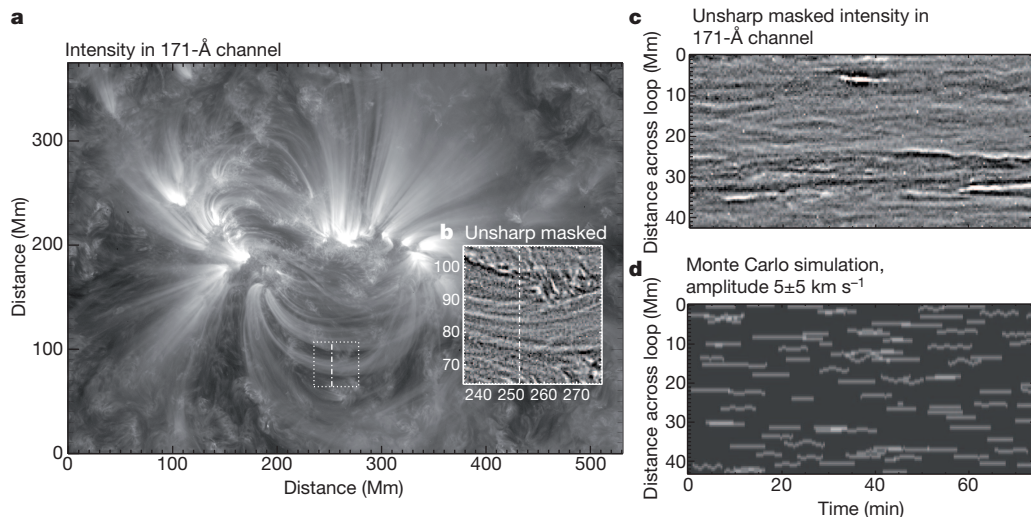
**Figure 1 | Ubiquitous Alfvénic motion above the solar limb.** Space-time plots of SDO/AIA data, demonstrating the visibility of the ubiquitous transverse waves 34 Mm above the solar limb in a coronal hole. **a**, Radial distance–position angle map, showing intensity in the 304-Å channel; the location studied is shown by the dashed horizontal line between the two vertical dashed lines. The intensity images (**b**, 304-Å channel) and mean-subtracted intensity images

(**d**, 171-Å channel) show a number of oscillatory structures that are enhanced using unsharp masking (**c** and **e**, respectively). We highlight one oscillation as an example; it is shown enclosed in a red rectangle, and is compatible with propagation along the spicule (**f**, 304-Å channel), and with propagating coronal disturbance (**g**, 171-Å channel). A sine wave with a period of 180 s and an amplitude of  $24 \text{ km s}^{-1}$  is drawn on **f** and **g** as a visual aid for the reader.



**Figure 2 | Examining Alfvénic motion in coronal hole (top row) and quiet Sun (bottom row) regions.** **a–e**, SDO/AIA space-time plots of unsharp masked intensity in the 304-Å (**a**, **d**) and 171-Å (**b**, **e**) channels 15 Mm above the solar limb (Supplementary Movies 3 and 4). **c**, **f**, Monte Carlo simulations for Alfvénic waves with periods of 150–600 s and amplitudes of  $25 (\pm 5)$  (**c**) and  $20 (\pm 5)$  (**f**)  $\text{km s}^{-1}$ . These simple simulations indicate that the spatio-temporal

superposition of many independent bright features carrying Alfvénic waves with random phases leads to poor visibility of the extrema, or ‘swings’ of the sinusoidal motion. This is because many of the sinusoidal swings are superimposed on top of features that do not show any apparent lateral motion (the polarization of the Alfvénic wave is along the line-of-sight).



**Figure 3 | Examining Alfvénic motion in an active region of the Sun.** An active region loop system was observed in the SDO/AIA 171-Å channel (a, taken on 25 April 2010, at 02:00 UT; Supplementary Movies 5–7). By isolating the coronal loops in the study region (boxed in a, shown magnified

and unsharp masked in b), we demonstrate the variation of the signal at one location (the dashed vertical line). From this we produce a space-time plot (c) for comparison with a Monte Carlo simulation (d; amplitude  $5 \pm 5 \text{ km s}^{-1}$ , period 100–400 s).

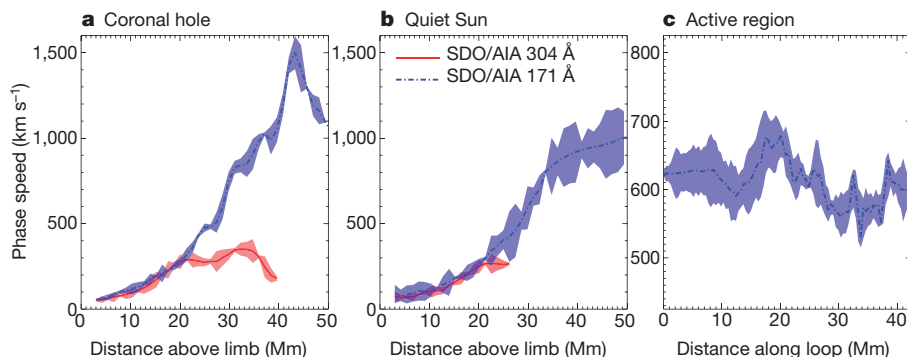
coronal hole and quiet corona ( $(5\text{--}10) \times 10^{-13} \text{ kg m}^{-3}$ ; Supplementary Information), we determine energy flux densities of the order of  $100\text{--}200 \text{ W m}^{-2}$  at a height of 15 Mm. The observed phase speeds and densities are compatible with the presence of magnetic fields of the order of 10 G.

The SDO/AIA observations of transition-region and coronal emission have permitted the measurement of the amplitude, period and phase speed of ubiquitous transverse waves. The estimated energy flux is sufficient to provide the energy necessary to drive the fast solar wind<sup>8</sup> and overcome the radiative losses of the quiet solar corona<sup>9</sup>.

The estimated energy flux of low-frequency Alfvénic waves in denser active region loops ( $\sim 100 \text{ W m}^{-2}$ ; Supplementary Information) is not sufficient to provide the entire  $2,000 \text{ W m}^{-2}$  required to power the active corona<sup>9</sup>. However, we warn that the phase speed of the disturbances seen in the active corona may be significantly under-reported, because bidirectional waves on the loops<sup>17</sup> can reduce the amplitude and phase speed. Further, we believe that instrumentation of higher spatial and temporal resolution than that of the SDO/AIA are required to study the full spectrum of wave energy (including higher frequencies) in active regions, because of the smaller length scales and shorter period

waves that are likely to exist there (Fig. 3). Therefore, the effect of Alfvénic waves on the energy balance of the active solar corona remains unclear, although our observations suggest that low-frequency Alfvénic waves do not carry enough energy to significantly affect active region heating.

Our results (for the quiet corona and coronal holes) suggest that a ‘two-stage’ process may be at play in the quiescent solar atmosphere. The first stage is an initial heating and injection of plasma from the lower atmosphere. This is supplemented by the second stage, which involves the dissipation of the Alfvénic waves; this dissipation can sustain the temperatures ( $\sim 10^6 \text{ K}$ ) of the coronal material in the quiet Sun, or accelerate the solar wind in a coronal hole. Such a secondary energy source in the magnetically closed regions of the corona can explain the gentle, steady evaporation of material from the upper chromosphere and transition region that is driven by downward thermal conduction<sup>18</sup>, as well as the stark spectroscopic contrast between the two<sup>19,20</sup>. Although caution is clearly necessary, the scheme we propose may not be too far-fetched: recent observational evidence of dissipation of Alfvénic waves in the quiet corona<sup>17</sup>, and many models, suggest that the transfer of energy from low-frequency Alfvénic waves



**Figure 4 | Determining the phase speed of the Alfvénic motions.** By cross-correlating the space-time plots at different heights above the limb, we can measure the phase speed of the observed disturbances and their error computed from an ensemble of cross-correlations about that height (dashed regions; Supplementary Information). Data are shown for coronal holes (a), quiet Sun (b) and along a coronal loop complex in an active region (c). In a and b we show results from both SDO/AIA channels (304 Å and 171 Å; c shows 171-Å results only. We see that the profiles of coronal hole and quiet Sun phase speed

determined from transition region emission rise to  $\sim 250 \text{ km s}^{-1}$  at a distance of 20 Mm above the solar limb, which is consistent with chromospheric measurements<sup>10</sup>. The continued increase to  $\sim 1,000 \text{ km s}^{-1}$  at a distance 40 Mm above the limb is consistent with previous coronal phase speed determinations<sup>7</sup>. The phase speed determined for the active region studied is of the order of  $600 \text{ km s}^{-1}$ , and a small variance ( $\sim 50 \text{ km s}^{-1}$ ) is observed along the loop structures.



to the plasma is essential to drive the fast solar wind to the high velocities observed at 1 AU (astronomical unit) from the Sun<sup>3–5,21–23</sup>. The challenge remains to understand how, and where, these waves are generated and dissipated in the solar atmosphere, and how that dissipation delivers energy to the ions and electrons that comprise the coronal plasma and solar wind<sup>6,23</sup>.

Received 16 February; accepted 25 May 2011.

1. Belcher, J. W. & Olbert, S. Stellar winds driven by Alfvén waves. *Astrophys. J.* **200**, 369–382 (1975).
2. Axford, W. I. *et al.* Acceleration of the high speed solar wind in coronal holes. *Space Sci. Rev.* **87**, 25–41 (1999).
3. Matthaeus, W. H., Zank, G. P., Oughton, S., Mullan, D. J. & Dmitruk, P. Coronal heating by magnetohydrodynamic turbulence driven by reflected low-frequency waves. *Astrophys. J.* **523**, L93–L97 (1999).
4. Cranmer, S. R., van Ballegoijen, A. A. & Edgar, R. J. Self-consistent coronal heating and solar wind acceleration from anisotropic magnetohydrodynamic turbulence. *Astrophys. J.* **171** (Suppl.), 520–551 (2007).
5. Suzuki, T. K. & Inutsuka, S.-i. Solar winds driven by nonlinear low-frequency Alfvén waves from the photosphere: parametric study for fast/slow winds and disappearance of solar winds. *J. Geophys. Res.* **111**, A06101, doi:10.1029/2005JA011502 (2006).
6. Verdini, A., Velli, M., Matthaeus, W. H., Oughton, S. & Dmitruk, P. A turbulence-driven model for heating and acceleration of the fast wind in coronal holes. *Astrophys. J.* **708**, L116–L120 (2010).
7. Tomczyk, S. *et al.* Observations of Alfvén waves in the quiet solar corona. *Science* **317**, 1192–1196 (2007).
8. Hansteen, V. H. & Leer, E. Coronal heating, densities, and temperatures and solar wind acceleration. *J. Geophys. Res.* **100**, 21577–21593 (1995).
9. Withbroe, G. L. & Noyes, R. W. Mass and energy flow in the solar chromosphere and corona. *Annu. Rev. Astron. Astrophys.* **15**, 363–387 (1977).
10. De Pontieu, B. *et al.* Chromospheric Alfvénic waves strong enough to power the solar wind. *Science* **318**, 1574–1577 (2007).
11. Tsuneta, S. *et al.* The solar optical telescope for the Hinode mission: an overview. *Sol. Phys.* **249**, 167–196 (2008).
12. Beckers, J. M. Solar spicules. *Sol. Phys.* **3**, 367–433 (1968).
13. De Pontieu, B. *et al.* A tale of two spicules: the impact of spicules on the magnetic chromosphere. *Publ. Astron. Soc. Jpn* **59**, 655–662 (2007).
14. Lemen, J. R. *et al.* The Atmospheric Imaging Assembly on the Solar Dynamics Observatory. *Sol. Phys.* (submitted).
15. De Pontieu, B. *et al.* The origin of hot coronal plasma. *Science* **331**, 55–58 (2011).
16. De Pontieu, B., McIntosh, S. W., Hansteen, V. H. & Schrijver, C. J. Observing the roots of solar coronal heating—in the chromosphere. *Astrophys. J.* **701**, L1–L6 (2009).
17. Tomczyk, S. & McIntosh, S. W. Time-distance seismology of the solar corona with CoMP. *Astrophys. J.* **697**, 1384–1391 (2009).
18. Patsourakos, S. & Klimchuk, J. A. Nonthermal spectral line broadening and the nanoflare model. *Astrophys. J.* **647**, 1452–1465 (2006).
19. McIntosh, S. W. *et al.* Observations supporting the role of magnetoconvection in energy supply to the quiescent solar atmosphere. *Astrophys. J.* **654**, 650–664 (2007).
20. McIntosh, S. W., Leamon, R. J. & De Pontieu, B. The spectroscopic footprint of the fast solar wind. *Astrophys. J.* **727**, 7 (2011).
21. Hollweg, J. V. Alfvén waves in a two-fluid model of the solar wind. *Astrophys. J.* **181**, 547–566 (1973).
22. Chandran, B. D. G. Alfvén-wave turbulence and perpendicular ion temperatures in coronal holes. *Astrophys. J.* **720**, 548–554 (2010).
23. Cranmer, S. R. & van Ballegoijen, A. A. Can the solar wind be driven by magnetic reconnection in the Sun's magnetic carpet? *Astrophys. J.* **720**, 824–847 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** SDO is the first mission of NASA's Living With a Star Program. NCAR is sponsored by the NSF.

**Author Contributions** S.W.M. (with B.D.P. and M.C.) performed all image processing and analysis of observations. S.W.M. and M.C. calculated phase speeds. P.B. (with B.D.P. and V.H.H.) designed the special observing sequences. B.D.P. co-aligned the data, performed Monte Carlo simulations (with V.H.H.) and provided density estimates. M.G. assisted with the identification of the wave mode. S.W.M. and B.D.P. wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to S.W.M. ([msscott@ucar.edu](mailto:msscott@ucar.edu)).

# Earth's Trojan asteroid

Martin Connors<sup>1,2</sup>, Paul Wiegert<sup>3</sup> & Christian Veillet<sup>4</sup>

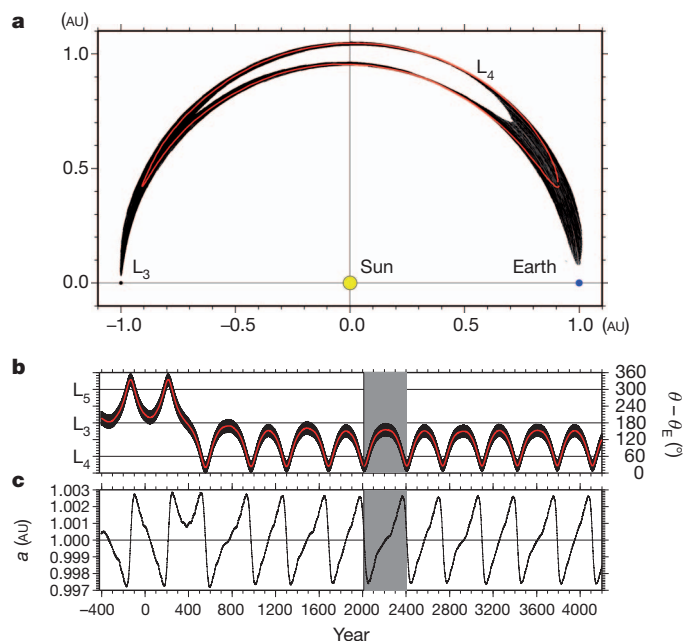
It was realized in 1772 that small bodies can stably share the same orbit as a planet if they remain near 'triangular points' 60° ahead of or behind it in the orbit<sup>1</sup>. Such 'Trojan asteroids' have been found co-orbiting with Jupiter<sup>2</sup>, Mars<sup>3</sup> and Neptune<sup>4</sup>. They have not hitherto been found associated with Earth, where the viewing geometry poses difficulties for their detection<sup>5</sup>, although other kinds of co-orbital asteroid (horseshoe orbiters<sup>6</sup> and quasi-satellites<sup>7</sup>) have been observed<sup>8</sup>. Here we report an archival search of infrared data for possible Earth Trojans, producing the candidate 2010 TK<sub>7</sub>. We subsequently made optical observations which established that 2010 TK<sub>7</sub> is a Trojan companion of Earth, librating around the leading Lagrange triangular point, L<sub>4</sub>. Its orbit is stable over at least ten thousand years.

The existence of Trojan asteroids of other planets raises the question of whether such companions could exist for Earth. Despite studies showing that such bodies could be relatively stable<sup>9</sup> and may wander relatively far from the Lagrange points<sup>5</sup>, they would dwell mostly in the daylight sky as seen from Earth, making detection difficult. Indeed, they hitherto have not been observed<sup>10,11</sup>. The launch of the Wide-field Infrared Survey Explorer (WISE) by NASA in 2009<sup>12</sup> provided improved viewing circumstances that made possible new detections of over 500 near-Earth objects<sup>13</sup>. WISE searched large areas of sky always 90° from the Sun, with high efficiency for asteroidal bodies and good astrometric accuracy. Examining WISE discoveries in the expectation that Earth co-orbital objects, possibly including a Trojan, could be found, resulted in two promising candidates, 2010 SO<sub>16</sub> and 2010 TK<sub>7</sub>. Both are larger than most co-orbital objects, being several hundred metres in diameter, and 2010 SO<sub>16</sub> is a horseshoe orbiter<sup>14</sup>. We identified 2010 TK<sub>7</sub> as probably being an Earth Trojan, on the basis of positions measured over a six-day arc in late 2010. Observations made at the University of Hawaii (D. Tholen, personal communication) and the Canada–France–Hawaii Telescope<sup>15</sup> in April 2011, after the object had for months been in an unfavourable position as seen from Earth, so greatly improved the knowledge of its orbit that we can state with certainty that 2010 TK<sub>7</sub> is an Earth Trojan.

The 'tadpole' motion of 2010 TK<sub>7</sub>, which is characteristic of Trojan asteroids, is shown in Fig. 1 in the frame co-rotating with Earth (see Supplementary Information for three-dimensional depictions of the motion). The 1-yr-averaged curve shows the centre of motion librating about L<sub>4</sub>, the Lagrange point 60° ahead of Earth. The period of this motion is at present 395 yr. Superposed on this is an annual motion or epicycle<sup>2,16,17</sup> (not shown for clarity). This mode of display emphasizes the longitudinal motion despite the enhanced radial scale: the asteroid's mean position drifts along the red line, from the 'head' of the tadpole, near Earth, to the far 'tail', where it is nearly on the opposite side of the Sun from the Earth. The relatively large eccentricity, of  $e = 0.191$ , results in an annual heliocentric radial motion between roughly 0.81 and 1.19 AU. The inclination of 2010 TK<sub>7</sub> is about  $i = 20.9^\circ$ , so there is significant motion perpendicular to Earth's orbital plane. The asteroid's eccentricity and inclination produce a large epicycle, which is responsible for the visibility of the object at the solar elongation of 90°, as observed by WISE; and it is now at the near-Earth end of the tadpole. In the present epoch, the longitude remains in the sector of L<sub>4</sub>, trapped

between Earth and L<sub>3</sub>. Interaction with Earth at the near-Earth end of the tadpole results in a rapid decrease in the object's semimajor axis,  $a$ , making it increase its angular speed (Kepler's third law) and outpace Earth. This is currently taking place. Slow resonant interaction at the other parts of the tadpole increases  $a$ , making the object slow gradually such that it again approaches Earth. In the current cycle, this will take place in the years AD 2050–2350, approximately. Repetition of this cycle leads to a sawtooth pattern in  $a$  as a function of time (Fig. 1c).

The present motion of 2010 TK<sub>7</sub> is well established, but there are inherent limits on our ability to compute orbits into the past or future. Chaos limits the accuracy of computations of this asteroid's position over timescales<sup>18</sup> greater than about 250 yr. However, we can still discuss the basic nature of its orbit with confidence by computing the motion of many 'dynamical clones' whose orbital parameters vary<sup>7</sup> within the limits set by observations. Approximately 1,800 yr in the past, and more than 5,000 yr in the future, the 100 clone orbits we



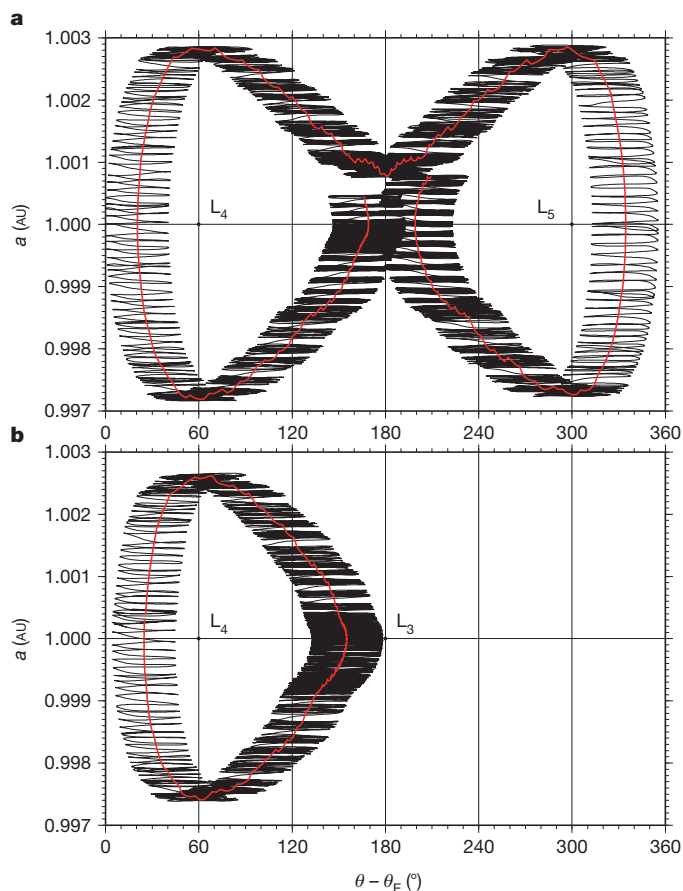
**Figure 1 | Orbital parameters of asteroid 2010 TK<sub>7</sub>.** **a**, Path over one Trojan libration from AD 2010 to 2405 in the co-rotating reference frame. In this frame, Earth is stationary and the average position of the asteroid librates about L<sub>4</sub> in a 'tadpole' orbit. Both Earth and the asteroid revolve about the Sun with periods close to 1 yr, and slow changes in their relative positions are best seen in the co-rotating frame. The difference between the asteroid's semimajor axis,  $a$ , and a circle of radius 1 AU (an astronomical unit (AU) is the Earth–Sun distance) is multiplied by a factor of 20 for clarity, and Earth and the Sun are not shown to scale. Black lines indicate  $a$  and longitude relative to Earth daily; the red curve shows the annual average. **b**, Longitude relative to Earth,  $\theta - \theta_E$ , over the period 420 BC to AD 4200. A 'jump' from L<sub>5</sub> libration to the present L<sub>4</sub> libration took place in around AD 400. Black and red lines indicate daily and averaged values, as in **a**. The grey band is the period of the present libration. **c**, Semimajor axis daily values. Initial conditions (best orbital solution) are given in Table 1.

<sup>1</sup>Athabasca University, 1 University Drive, Athabasca, Alberta T9S 3A3, Canada. <sup>2</sup>Department of Earth and Space Sciences, UCLA, Los Angeles, California 90095, USA. <sup>3</sup>Department of Physics and Astronomy, The University of Western Ontario, London, Ontario N6A 3K7, Canada. <sup>4</sup>Canada–France–Hawaii Telescope, Kamuela, Hawaii 96743, USA.



computed diverged sufficiently that we must say that the asteroid's precise behaviour cannot be predicted with certainty outside that  $\sim 7,000$ -yr span. The range of behaviour shown by the clones and, thus, possible for the real object includes making transitions to horseshoe modes and 'jumping' between Lagrange points. Short-term unstable libration about  $L_3$ , the Lagrange point on the other side of the Sun from Earth, can occur as a result of the asteroid's large inclination. Such orbits were theorized as early as 1920<sup>17</sup>, but no real object had until now been suspected to enter them.

Jumping from one Lagrange point to the other is a behaviour previously attributed to the Jupiter Trojan 1868 Thersites<sup>19</sup>, and was found in about half of the clone orbits. Here, the large eccentricity leads to longitudinal excursions (Kepler's second law), including when near  $L_3$ . In Fig. 2, these are shown to have allowed (in about AD 500) a rapid transition of 2010 TK<sub>7</sub> from  $L_5$  to the present  $L_4$  libration. The libration now remains only in the sector of  $L_4$  and is relatively stable, in a classic<sup>16</sup> Trojan pattern, although of large amplitude.



**Figure 2 | Semimajor axis versus relative longitude for 2010 TK<sub>7</sub>.**  
**a**, Libration during the period AD 1–800, featuring a 'jump' from libration initially about  $L_5$  (right) to the present libration around  $L_4$ . As in Fig. 1, black lines indicate daily values and red lines indicate the annual averages. When the asteroid is near a relative longitude,  $\theta - \theta_E$ , of about  $180^\circ$ , the annual excursions in relative longitude can cause it to cross  $L_3$ . This crossing can trigger a rapid transition or 'jump' between librational modes. Clone studies show that the chaotic behaviour of the asteroid is due mainly to a great sensitivity to non-resonant perturbations when near  $L_3$ . Libration about  $L_5$  results in an average longitude  $120^\circ$  different from that for libration about  $L_4$ . Such a large change resulting from small perturbations (when near  $L_3$ ) is characteristic of chaos.  
**b**, Present (AD 2010–2405) libration about  $L_4$ . The location of  $L_3$  is shown for reference but the relative longitude in the era after AD 800 does not cross it, resulting in the current stability of the orbit. The apparent banding in both panels is due to changes in semimajor axis and has a predominant period of roughly 12 yr; therefore, it is probably mainly caused by Jupiter perturbations.

**Table 1 | Heliocentric orbital elements of 2010 TK<sub>7</sub>**

Epoch	JD 2455600.5
Semimajor axis, $a$	1.0004078 AU
Eccentricity, $e$	0.1908177
Inclination, $i$	$20.87984^\circ$
Argument of perihelion	$45.86009^\circ$
Longitude of ascending node	$96.54190^\circ$
Mean anomaly	$20.30069^\circ$

Results in the figures were obtained using these initial conditions in the Mercury integrator<sup>24</sup> (verified in the near-present using the JPL Horizons system<sup>25</sup>). The RADAU<sup>26</sup> option was used with 1-d spacing for the eight planets, Pluto and the Earth–Moon barycentre. Clone studies included the eight planets and the Moon<sup>27</sup>, with variations<sup>7</sup> among the orbital elements of the order of the last significant digit shown. The Julian date (JD) shown corresponds to 0:00 UT on 8 February 2011.

Chaotic effects have a large role in the behaviour of this asteroid. Its sensitivity to small influences when in the vicinity of  $L_3$  allows the range of outcomes seen among the clones. The overall Trojan behaviour is dictated by 1:1 orbital resonance with Earth, but non-resonant effects of Jupiter are 80 times stronger than those of Earth when Jupiter is at the same celestial longitude as  $L_3$ . Such influences, demonstrated by the 'banding' seen in Fig. 2, alter the asteroid's chaotic behaviour. Many clone orbits make repeated transitions between the Lagrange points, such that the chaos can be stable<sup>20</sup>, with  $L_4$  and  $L_5$  each defining permitted regions of phase space. Knowledge of the orbit will improve as it is observed over the years, but its chaotic nature dictates that dynamics-based discussions of the origin and fate of 2010 TK<sub>7</sub>, and its relationship to other bodies, will necessarily remain statistical in nature.

Earth Trojan asteroids have been proposed as natural candidates for spacecraft rendezvous missions<sup>21</sup>. However, the large inclination of 2010 TK<sub>7</sub> results in a  $\Delta v$  of  $9.4 \text{ km s}^{-1}$  being required, whereas other near-Earth asteroids have values of  $\Delta v$  less than  $4 \text{ km s}^{-1}$ . The reported absolute magnitude, 20.7 mag, puts the diameter of 2010 TK<sub>7</sub> at 300 m with an assumed albedo of 0.1 (ref. 22), which makes it relatively large among the near-Earth asteroid population. No spectral or colour information is as yet available to determine whether the asteroid is in any other way unusual.

Received 11 April; accepted 27 May 2011.

- Wilson, C. in *Planetary Astronomy from the Renaissance to the Rise of Astrophysics. Part B: The Eighteenth and Nineteenth Centuries* (eds Taton, R. & Wilson, C.) 108–130 (Cambridge, 1995).
- Stacey, R. G. & Connors, M. A century survey of orbits of co-orbitals of Jupiter. *Planet. Space Sci.* **56**, 358–367 (2008).
- Connors, M. *et al.* A survey of orbits of co-orbitals of Mars. *Planet. Space Sci.* **53**, 617–624 (2005).
- Almeida, A. J. C., Peixinho, N. & Correia, A. C. M. Neptune Trojans and Plutinos: colors, sizes, dynamics, and their possible collisions. *Astron. Astrophys.* **508**, 1021–1030 (2009).
- Wiegert, P., Innanen, K. A. & Mikkola, S. Earth Trojan asteroids: a study in support of observational searches. *Icarus* **145**, 33–43 (2000).
- Wiegert, P., Innanen, K. A. & Mikkola, S. An asteroidal companion to the Earth. *Nature* **387**, 685–686 (1997).
- Brasser, R. *et al.* Transient co-orbital asteroids. *Icarus* **171**, 102–109 (2004).
- Connors, M. *et al.* Discovery of an asteroid and quasi-satellite in an Earth-like horseshoe orbit. *Meteor. Planet. Sci.* **37**, 1435–1441 (2002).
- Mikkola, S. & Innanen, K. A. Studies in solar system dynamics. II. The stability of Earth's Trojans. *Astron. J.* **100**, 290–293 (1990).
- Whitely, R. J. & Tholen, D. J. A CCD search for Lagrangian asteroids of the Earth–Sun system. *Icarus* **136**, 154–167 (1998).
- Connors, M. *et al.* Initial results of a survey of Earth's  $L_4$  point for possible Earth Trojan asteroids. *Bull. Am. Astron. Soc.* **32**, 1019 (2000).
- Wright, E. L. *et al.* The Wide-field Infrared Survey Explorer (WISE): mission description and initial on-orbit performance. *Astron. J.* **140**, 1868–1881 (2010).
- Mainzer, A. *et al.* Preliminary results from NEOWISE: an enhancement to the Wide-Field Infrared Survey Explorer for solar system science. *Astrophys. J.* **731**, 53 (2011).
- Christou, A. A. & Asher, D. J. A long-lived horseshoe companion to the Earth. *Mon. Not. R. Astron. Soc.* advance online publication, doi:10.1111/j.1365-2966.2011.18595.x (12 April 2011).
- Boulaide, O. *et al.* Development of MegaCam, the next-generation wide-field imaging camera for the 3.6-m Canada-France-Hawaii Telescope. *Proc. SPIE* **2008**, 657–668 (2000).
- Murray, C. D. & Dermott, S. F. *Solar System Dynamics* 63–110 (Cambridge Univ. Press, 1999).
- Moulton, F. R. *Periodic Orbits* 299–324, 151–198 (Carnegie Institute, 1920).
- Whipple, A. L. Lyapunov times of the inner asteroids. *Icarus* **115**, 347–353 (1995).

19. Tsiganis, K., Dvorak, R. & Pilat-Lohinger, E. Thersites: a 'jumping' Trojan? *Astron. Astrophys.* **354**, 1091–1100 (2000).
20. Milani, A. & Nobili, A. M. An example of stable chaos in the Solar System. *Nature* **357**, 569–571 (1992).
21. Stacey, R. G. & Connors, M. Delta-v requirements for earth co-orbital rendezvous missions. *Planet. Space Sci.* **57**, 822–829 (2009).
22. Bowell, E. *et al.* in *Asteroids II* (eds Binzel, R., Gehrels, T. & Matthews, M.) 524–556 (Univ. Arizona Press, 1989).
23. MPEC 2011-G37: Daily Orbit Update (2011 Apr. 5 UT). *IAU Minor Planet Center* (<http://www.minorplanetcenter.org/mpec/K11/K11G37.html>) (2011).
24. Chambers, J. E. A hybrid symplectic integrator that permits close encounters between massive bodies. *Mon. Not. R. Astron. Soc.* **304**, 793–799 (1999).
25. Giogini, J. D. *et al.* JPL's online solar system data service. *Bull. Am. Astron. Soc.* **28**, 1158 (1996).
26. Everhart, E. in *Dynamics of Comets: Their Origin and Evolution* (eds Carusi, A. & Valsecchi, G. B.) 185–202 (Kluwer, 1985).
27. Standish, E. M. *Planetary and Lunar Ephemerides DE406/LE406*. Tech. Report IOM 312.F-98-048 (NASA Jet Propulsion Laboratory, 1998).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the WISE team and JPL and NEODyS (University of Pisa) data services. Support came from Canada's Natural Sciences and Engineering Research Council and Research Chairs. Problems with some 2010 TK<sub>7</sub> positions reported in ref. 23 were pointed out by T. Spahr and D. Tholen. We are grateful to them for data reductions provided, and to D. Tholen, M. Micheli and G. T. Elliot for observations made in support of this study.

**Author Contributions** The authors contributed equally to this work. M.C. and P.W. concentrated on dynamical calculations, and C.V. concentrated on observations and associated data reduction.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.C. ([martinc@athabascau.ca](mailto:martinc@athabascau.ca)).



# Two-photon laser spectroscopy of antiprotonic helium and the antiproton-to-electron mass ratio

Masaki Hori<sup>1,2</sup>, Anna Sótér<sup>1</sup>, Daniel Barna<sup>2,3</sup>, Andreas Dax<sup>2</sup>, Ryugo Hayano<sup>2</sup>, Susanne Friedreich<sup>4</sup>, Bertalan Juhász<sup>4</sup>, Thomas Pask<sup>4</sup>, Eberhard Widmann<sup>4</sup>, Dezső Horváth<sup>3,5</sup>, Luca Venturelli<sup>6</sup> & Nicola Zurlo<sup>6</sup>

Physical laws are believed to be invariant under the combined transformations of charge, parity and time reversal (CPT symmetry<sup>1</sup>). This implies that an antimatter particle has exactly the same mass and absolute value of charge as its particle counterpart. Metastable antiprotonic helium ( $\bar{p}\text{He}^+$ ) is a three-body atom<sup>2</sup> consisting of a normal helium nucleus, an electron in its ground state and an antiproton ( $\bar{p}$ ) occupying a Rydberg state with high principal and angular momentum quantum numbers, respectively  $n$  and  $l$ , such that  $n \approx l + 1 \approx 38$ . These atoms are amenable to precision laser spectroscopy, the results of which can in principle be used to determine the antiproton-to-electron mass ratio and to constrain the equality between the antiproton and proton charges and masses. Here we report two-photon spectroscopy of antiprotonic helium, in which  $\bar{p}^3\text{He}^+$  and  $\bar{p}^4\text{He}^+$  isotopes are irradiated by two counter-propagating laser beams. This excites nonlinear, two-photon transitions of the antiproton of the type  $(n, l) \rightarrow (n-2, l-2)$  at deep-ultraviolet wavelengths ( $\lambda = 139.8, 193.0$  and  $197.0$  nm), which partly cancel the Doppler broadening of the laser resonance caused by the thermal motion of the atoms. The resulting narrow spectral lines allowed us to measure three transition frequencies with fractional precisions of 2.3–5 parts in  $10^9$ . By comparing the results with three-body quantum electrodynamics calculations, we derived an antiproton-to-electron mass ratio of  $1,836.1526736(23)$ , where the parenthetical error represents one standard deviation. This agrees with the proton-to-electron value known to a similar precision.

Antiprotonic atoms (denoted  $\bar{p}X^+$ ) can be readily synthesized in a given element  $X$  by replacing the atomic electrons with a negatively charged antiproton<sup>2</sup>. The substitution takes place spontaneously when antiprotons are brought to rest in the substance concerned. However, these exotic atoms are usually destroyed within picoseconds by electromagnetic cascade mechanisms that result in the annihilation of the antiprotons with the nucleus of  $X$ . The  $\bar{p}\text{He}^+$  atom alone has microsecond-scale lifetimes even in dense helium targets. The extreme longevity is due to the fact that the antiprotons trapped in the  $n \approx l + 1 \approx 38$  Rydberg states have almost no overlap with the nucleus, and furthermore cannot easily de-excite by Auger emission of the electron owing to its large binding energy ( $I \approx 25$  eV) and the large multiplicity ( $\Delta l$ ) of the necessary transition. The electron protects the antiproton during collisions with other helium atoms, making  $\bar{p}\text{He}^+$  amenable to laser spectroscopy.

The energy levels of  $\bar{p}\text{He}^+$  have been calculated by three-body quantum electrodynamics (QED) calculations (ref. 3 and V. I. Korobov, personal communication) to precisions of  $1 \times 10^{-9}$ . The calculated values now include nuclear size effects, and relativistic and radiative recoil corrections up to order  $m\alpha^6$ , where  $m$  and  $\alpha$  respectively denote the electron mass and the fine structure constant. The fractional measurement precision of single-photon laser spectroscopy experiments<sup>4,5</sup> of  $\bar{p}\text{He}^+$ , however, has always been limited to  $10^{-7}$ – $10^{-8}$  by the Doppler broadening effect. As in normal atoms,

the thermal motion of  $\bar{p}\text{He}^+$  at temperature  $T$  strongly broadens the measured widths of the laser resonances, by  $\sim \nu \sqrt{8k_B T \log(2)/Mc^2}$ , where  $\nu$  denotes the transition frequency,  $k_B$  the Boltzmann constant,  $M$  the atom's mass and  $c$  the speed of light.

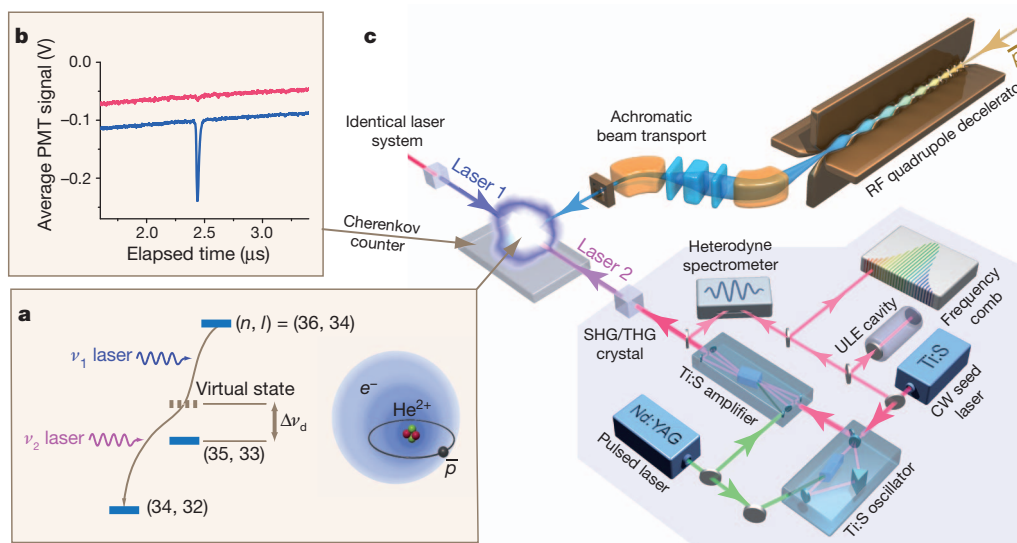
One way to reach precisions beyond this Doppler limit is provided by two-photon spectroscopy. For example, the  $1s$ – $2s$  transition frequency in atomic hydrogen (H) has been measured to a precision of  $10^{-14}$  using two counter-propagating laser beams, each with a frequency of half the  $1s$ – $2s$  energy interval. This arrangement cancels the Doppler broadening to first order<sup>6</sup>. A similar experiment has been proposed for antihydrogen ( $\bar{\text{H}}$ ) atoms<sup>7,8</sup>, a sample of which was recently confined in a magnetic trap<sup>8</sup>. However, it is normally difficult to apply this to  $\bar{p}\text{He}^+$  because of the small probabilities involved in the nonlinear transitions of the massive antiproton. Indeed, calculations<sup>9</sup> show that gigawatt-scale laser powers would be needed to excite them against annihilation within the atom's lifetime.

Nevertheless, we have been able to induce two-photon antiproton transitions of the type  $(n, l) \rightarrow (n-2, l-2)$  (Fig. 1a), by using the fact that this small transition probability can be enhanced by factor  $>10^5$  if the counter-propagating beams have frequencies,  $\nu_1$  and  $\nu_2$ , such that the virtual intermediate state of the two-photon transition lies within  $\Delta\nu_d \approx 10$  GHz of a real state<sup>10</sup>,  $(n-1, l-1)$ . At resonance between the atom and the laser beams, the antiprotons are directly transferred between the parent and daughter states via the nonlinear transition, leaving the population in  $(n-1, l-1)$  unaffected. The first-order Doppler width is then reduced by a factor of  $|\nu_1 - \nu_2|/(\nu_1 + \nu_2)$ .

In these experiments, transitions were selected for laser excitation between pairs of states with microsecond- and nanosecond-scale lifetimes against Auger emission of the electron. At the two-photon resonance, Auger decay left a  $\bar{p}\text{He}^{2+}$  ion behind. No longer protected by the electron in the way described above, the ion was rapidly destroyed in Stark collisions with other helium atoms. Charged pions emerged from the resulting antiproton annihilations, passed through an acrylic sheet and produced Cherenkov radiation, which was detected by photomultipliers. The two-photon resonance condition between the counter-propagating lasers beams and the atom was thus revealed as a sharp spike in the annihilation rate (Fig. 1b).

Even under these conditions of enhanced transition probability, megawatt-scale laser pulses of high spectral purity and low phase noise are needed to excite these two-photon transitions and avoid rapid dephasing of the transition amplitude<sup>10</sup>. For this (Fig. 1c), we developed two sets of Ti:sapphire lasers with pulse lengths of 30–100 ns and among the smallest linewidths reported so far<sup>11</sup> ( $\sim 6$  MHz). They were based on continuous-wave lasers of wavelength 728–940 nm whose frequencies were measured to a precision of  $<1 \times 10^{-10}$  using a femtosecond optical comb<sup>12</sup> locked to a Global-Positioning-System-disciplined quartz oscillator. This continuous-wave seed beam was pulse-amplified to the necessary 1-MW peak power using a Ti:sapphire oscillator and amplifier. Spurious modulations in the pulsed laser frequency, or

<sup>1</sup>Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D85748 Garching, Germany. <sup>2</sup>Department of Physics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. <sup>3</sup>KFKI Research Institute for Particle and Nuclear Physics, H-1525 Budapest, Hungary. <sup>4</sup>Stefan Meyer Institut für Subatomare Physik, Boltzmanngasse 3, Vienna 1090, Austria. <sup>5</sup>Institute of Nuclear Research, H-4001 Debrecen, Hungary. <sup>6</sup>Dipartimento di Chimica e Fisica per l'Ingegneria e per i Materiali, Università di Brescia & Istituto Nazionale di Fisica Nucleare, Gruppo Collegato di Brescia, I-25133 Brescia, Italy.



**Figure 1 | Energy levels, Cherenkov detector signals and experimental layout for two-photon spectroscopy of  $\bar{p}\text{He}^+$ .** **a**, Two counter-propagating laser beams induced the two-photon transition  $(n, l) = (36, 34) \rightarrow (34, 32)$  in  $\bar{p}\text{He}^+$  via a virtual intermediate state of the antiproton tuned close to the real state  $(35, 33)$ . **b**, Cherenkov detectors revealed the annihilation of  $\bar{p}\text{He}^+$  following the nonlinear two-photon resonance induced at  $t = 2.4 \mu\text{s}$  (blue). When one of the lasers was detuned from resonance frequency by  $-500 \text{ MHz}$ , the two-photon signal abruptly disappeared (red). PMT, photomultiplier tube.

‘chirp’, induced during this amplification are an important source of systematic error<sup>5,13–15</sup> and were measured using a heterodyne spectrometer<sup>11</sup>. The precision of this laser system was verified<sup>11</sup> to be  $<1.4 \times 10^{-9}$  by measuring some two-photon transition frequencies in rubidium and caesium at respective wavelengths of 778 and 822 nm.

It was essential to use helium targets of low enough density for the relaxations caused by collisions between  $\bar{p}\text{He}^+$  and other helium atoms that could inhibit the two-photon transition to remain small. This implied the use of antiprotons of low enough energy to be stopped in such targets within the volume irradiated by the 2-cm-diameter laser beams. We used the CERN Antiproton Decelerator to produce 200-ns-long, pulsed beams of 5.3-MeV antiprotons (Fig. 1c). Every 100 s, we decelerated about  $7 \times 10^6$  antiprotons to  $\sim 70 \text{ keV}$  by allowing them to pass through a 3-m-long, radio-frequency quadrupole decelerator<sup>4</sup>. The beam was then transported by an achromatic, magnetic beamline to the target chamber filled with  $^4\text{He}$  or  $^3\text{He}$  gas at temperature  $T \approx 15 \text{ K}$  and pressure  $P = 0.8\text{--}3 \text{ mbar}$ . At a time 2–8  $\mu\text{s}$  after the resulting formation of  $\bar{p}\text{He}^+$ , two horizontally polarized laser beams of energy density  $\sim 1 \text{ mJ cm}^{-2}$  were simultaneously fired through the target in opposite directions perpendicular to the antiproton beam.

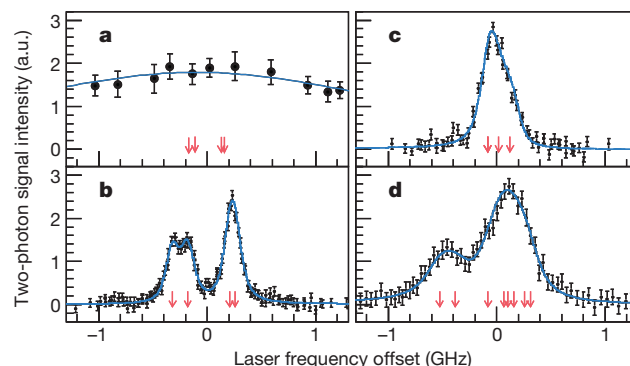
Figure 1b shows the Cherenkov signal (solid blue line) as a function of time elapsed since the arrival of antiproton pulses at the target, averaged over 30 pulses, which corresponds to  $\sim 10^7 \bar{p}\text{He}^+$  atoms. Laser beams of wavelengths  $c/v_1 = 417$  and  $c/v_2 = 372 \text{ nm}$  were tuned to the two-photon transition  $(36, 34) \rightarrow (34, 32)$  such that the virtual intermediate state lay  $\Delta v_d \approx 6 \text{ GHz}$  away from the real state  $(35, 33)$ . The above-mentioned annihilation spike corresponding to the two-photon transition can be seen at  $t = 2.4 \mu\text{s}$ . When the 417-nm laser alone was tuned off the two-photon resonance condition slightly (by 0.5 GHz; Fig. 1b, red line), the signal abruptly disappeared as expected. This indicates that the background from any Doppler-broadened, single-photon transitions is very small.

Figure 2b shows the resonance profile measured by detuning the laser of frequency  $v_2$  by  $\Delta v_d = -6 \text{ GHz}$  and scanning the laser of frequency  $v_1$  between  $-1$  and  $1 \text{ GHz}$  around the two-photon resonance defined by  $v_1 + v_2$ , which corresponds to a wavelength of  $\sim 197.0 \text{ nm}$ . The

**c**, The  $\bar{p}\text{He}^+$  atoms were synthesized by decelerating a beam of antiprotons using a radio-frequency quadrupole, and allowing them to stop in a cryogenic helium target. Two Ti:sapphire pulsed lasers whose optical frequencies were stabilized to a femtosecond frequency comb were used to carry out the spectroscopy. CW, continuous wave; RF, radio frequency; SHG, second-harmonic generation; THG, third-harmonic generation; ULE, ultralow expansion.

measured linewidth ( $\sim 200 \text{ MHz}$ ) represents the highest spectral resolution achieved so far for an antiprotonic atom, and is more than an order of magnitude smaller than the Doppler- and power-broadened profile of the corresponding single-photon resonance  $(36, 34) \rightarrow (35, 33)$  (Fig. 2a) measured under the same target and laser power conditions. This allows us to determine the atomic transition frequency with a correspondingly higher precision. The remaining width is caused by the hyperfine structure; the 3-ns Auger lifetime of the daughter state,  $(34, 32)$ ; and power broadening effects.

The two-peak structure with a frequency interval of 500 MHz arises from the dominant interaction between the electron spin and the orbital angular momentum of the antiproton. Each peak is a superposition of two hyperfine lines caused by a further interaction between the antiproton and electron spins. The asymmetric structure is reproduced by line shape calculations<sup>9</sup> (see below) and is due to the 25-MHz



**Figure 2 | Profiles of sub-Doppler two-photon resonances.** **a**, Doppler- and power-broadened profile of the single-photon resonance  $(36, 34) \rightarrow (35, 33)$  of  $\bar{p}\text{He}^+$ . **b**, Sub-Doppler two-photon profile of  $(36, 34) \rightarrow (34, 32)$  involving the same parent state. **c**, **d**, Profiles of  $(33, 32) \rightarrow (31, 30)$  of  $\bar{p}\text{He}^+$  (**c**) and  $(35, 33) \rightarrow (33, 31)$  of  $\bar{p}\text{He}^+$  (**d**). Black filled circles indicate experimental data points with 1-s.d. error bars, blue lines are best fits of theoretical line profiles (see text) and partly overlapping arrows indicate positions of the hyperfine lines. a.u., arbitrary units.



**Table 1 | Spin-averaged transition frequencies of  $\bar{p}\text{He}^+$** 

Isotope	Transition ( $n, l \rightarrow n-2, l-2$ )	Transition frequency (MHz)	
		Experiment	Theory
$\bar{p}^4\text{He}^+$	(36, 34) $\rightarrow$ (34, 32)	1,522,107,062(4)(3)(2)	1,522,107,058.9(2.1)(0.3)
	(33, 32) $\rightarrow$ (31, 30)	2,145,054,858(5)(5)(2)	2,145,054,857.9(1.6)(0.3)
$\bar{p}^3\text{He}^+$	(35, 33) $\rightarrow$ (33, 31)	1,553,643,100(7)(7)(3)	1,553,643,100.7(2.2)(0.2)

Experimental values show respective total, statistical and systematic 1-s.d. errors in parentheses; theoretical values (ref. 3 and V. I. Korobov, personal communication) show respective uncertainties from uncalculated QED terms and numerical errors in parentheses.

spacing between the unresolved hyperfine lines ( $S_e, S^p$ ) = ( $\uparrow\uparrow$ )  $\rightarrow$  ( $\uparrow\uparrow$ ) and ( $\uparrow\downarrow$ )  $\rightarrow$  ( $\uparrow\downarrow$ ) being smaller than the 75-MHz spacing between ( $\uparrow\uparrow$ )  $\rightarrow$  ( $\downarrow\uparrow$ ) and ( $\downarrow\downarrow$ )  $\rightarrow$  ( $\downarrow\downarrow$ ).

We next detected the (33, 32)  $\rightarrow$  (31, 30) resonance at wavelength  $\lambda = 139.8$  nm with the lowest  $n$  values among the two-photon transitions, using lasers of wavelengths  $c/v_1 = 296$  and  $c/v_2 = 264$  nm (Fig. 2c). The small transition probability and antiproton population required that higher laser intensities,  $P > 2$  mJ cm $^{-2}$ , and small detunings,  $\Delta v_d \approx 3$  GHz, from state (32, 31) were needed. For this transition, the four hyperfine lines are much closer together, lying within a 200-MHz range. We also measured the  $\bar{p}^3\text{He}^+$  resonance (35, 33)  $\rightarrow$  (33, 31) at  $\lambda = 139.8$  nm (Fig. 2d) using lasers of wavelengths  $c/v_1 = 410$  and  $c/v_2 = 364$  nm. This profile contains eight partly overlapping hyperfine lines arising from the spin-spin interactions of the  $^3\text{He}$  nucleus, the electron and the antiproton.

We determined the spin-independent transition frequencies,  $\nu_{\text{exp}}$  (Table 1), by fitting each profile with a theoretical line shape<sup>9</sup> (Fig. 2, blue lines) that was determined by numerically solving the nonlinear rate equations of the two-photon process. This included taking into account all two-photon transitions between the  $2l+1 \approx 70$  substates, the transition rates, power broadening effects, thermal motion of the atoms, the spurious frequency modulation<sup>11</sup> in the laser pulse, the experimentally measured spatial and temporal profiles of the laser beam, and a.c. Stark effects<sup>9</sup>. The positions of the hyperfine lines were fixed to the theoretical values, which have a precision of  $<0.5$  MHz (ref. 16).

For the transition (36, 34)  $\rightarrow$  (34, 32) in  $\bar{p}^4\text{He}^+$  (Table 2), the statistical error,  $\sigma_{\text{stat}}$ , due to the finite number of atoms in the laser beam was estimated to be 3 MHz (all quoted errors are s.d.). We measured transitions at various target densities between  $1 \times 10^{18}$  and  $3 \times 10^{18}$  cm $^{-3}$ . Within this density range, no significant collisional shift was observable within the 3-MHz experimental error. This agrees with quantum chemistry calculations (ref. 17 and D. Bakalov *et al.*, personal communication) for which the predictions of 0.1–1-MHz-scale collisional shifts in the associated single-photon lines agreed with experimental results<sup>4,18</sup> to within  $\sim 20\%$ . Calculations show that

**Table 2 | Errors for transition ( $n, l$ ) = (36, 34)  $\rightarrow$  (34, 32) of  $\bar{p}^4\text{He}^+$** 

Datum	Error (MHz)
Experimental errors	
Statistical error, $\sigma_{\text{stat}}$	3
Collisional shift error	1
A.c. Stark shift error	0.5
Zeeman shift	$<0.5$
Frequency chirp error	0.8
Seed laser frequency calibration	$<0.1$
Hyperfine structure	$<0.5$
Line profile simulation	1
Total systematic error, $\sigma_{\text{sys}}$	1.8
Total experimental error, $\sigma_{\text{exp}}$	3.5
Theoretical uncertainties	
Uncertainties from uncalculated QED terms*	2.1
Numerical uncertainty in calculation*	0.3
Mass uncertainties*	$<0.1$
Charge radii uncertainties*	$<0.1$
Total theoretical uncertainty*, $\sigma_{\text{th}}$	2.1

Experimental errors and theoretical uncertainties are 1 s.d.

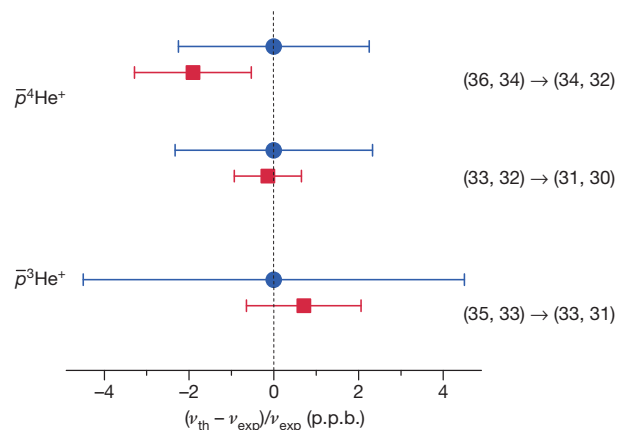
\*Ref. 3 and V. I. Korobov, personal communication.

magnetic Zeeman shifts are also small ( $<0.5$  MHz) for the Rydberg states under our experimental conditions. The frequency chirp of each laser pulse was recorded and corrected to a precision<sup>11</sup> of 0.8 MHz. We estimated the systematic error arising from the calculation of the fitting function<sup>9</sup> to be around 1 MHz.

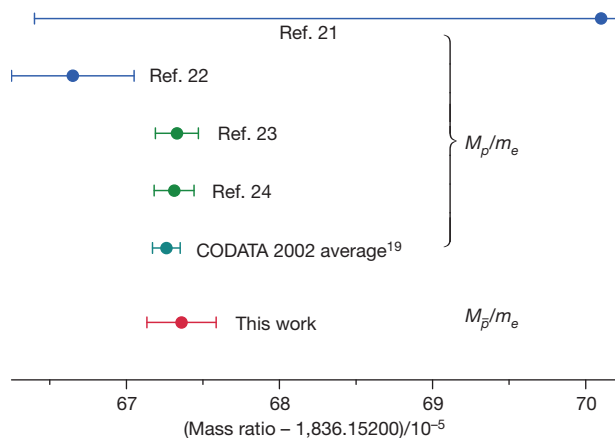
Laser fields can shift the frequencies of the two-photon transitions<sup>9</sup> by an amount proportional to  $(\Omega_1 - \Omega_2)/\Delta v_d$ , where  $\Omega_1$  and  $\Omega_2$  denote the Rabi frequencies of transitions between the parent and virtual intermediate states and, respectively, the daughter and intermediate states. We reduced this a.c. Stark shift to  $\leq 5$  MHz by carefully adjusting the intensities of the two laser beams such that  $\Omega_1 \approx \Omega_2$ . Remaining shifts were cancelled to a level of 0.5 MHz by systematically comparing<sup>9</sup> the resonance profiles measured alternately at positive and negative detunings,  $\pm \Delta v_d$ . The total experimental error,  $\sigma_{\text{exp}}$ , was obtained as the quadratic sum of all these errors. The larger error for the 193.0-nm  $\bar{p}^3\text{He}^+$  transitions is due to the larger number (eight) of hyperfine lines and the smaller signal intensity.

The experimental transition frequencies,  $\nu_{\text{exp}}$  (Fig. 3, filled circles with error bars in), agree with theoretical values,  $\nu_{\text{th}}$  (squares), to within  $(2-5) \times 10^{-9}$ . This agreement is a factor of five to ten times better than that obtained in previous single-photon experiments<sup>5</sup>. The calculation uses fundamental constants<sup>19</sup> compiled in CODATA 2002 including the  $^3\text{He}$ -to-electron and  $^4\text{He}$ -to-electron mass ratios, the Bohr radius and the Rydberg constant. To preserve the independence of this work, we avoided using the more recent CODATA 2006 (<http://physics.nist.gov/cuu/Constants/archive2006.html>) values, which include results from our previous experiments and three-body QED calculations on  $\bar{p}\text{He}^+$ . The charge radii of the  $^3\text{He}$  and  $^4\text{He}$  nuclei give corrections to  $\nu_{\text{th}}$  of 4–7 MHz, whereas the correction from the antiproton radius is much smaller (refs 3,20;  $\ll$  MHz) owing to the large  $l$  values of the states. The precision of  $\nu_{\text{th}}$  is mainly limited by the uncalculated radiative corrections of order  $m\alpha^8$  (Table 2).

When the antiproton-to-electron mass ratio,  $M_{\bar{p}}/m_e$ , in these calculations was changed by  $10^{-9}$ ,  $\nu_{\text{th}}$  changed by 2.3–2.8 MHz. By minimizing  $\sum_{\bar{p}} [\nu_{\text{th}}(M_{\bar{p}}/m_e) - \nu_{\text{exp}}]^2 / \sigma_{\text{stat}}^2$ , where the sum is over the three  $\bar{p}\text{He}^+$  frequencies, and considering the above systematic errors,  $\sigma_{\text{sys}}$ ,



**Figure 3 | Two-photon transition frequencies.** The experimental values ( $\nu_{\text{exp}}$ ; blue circles) for  $\bar{p}^4\text{He}^+$  and  $\bar{p}^3\text{He}^+$  agree with theoretical values ( $\nu_{\text{th}}$ ; red squares) to within fractional precisions of  $(2-5) \times 10^{-9}$ . Error bars, 1 s.d.; p.p.b., parts per  $10^9$ .



**Figure 4 | Antiproton-to-electron and proton-to-electron mass ratios.** The antiproton-to-electron mass ratio determined in this work agrees to within a fractional precision of  $<1.3$  p.p.b. with the proton-to-electron values measured in previous experiments<sup>21–24</sup> and the CODATA 2002 recommended value obtained by averaging them<sup>19</sup>. Error bars, 1 s.d.

we obtained the ratio  $M_{\bar{p}}/m_e = 1,836.1526736(23)$ , which yielded the best agreement between theoretical and experimental frequencies. The uncertainty,  $2.3 \times 10^{-6}$ , includes the statistical and systematic experimental contributions, respectively  $1.8 \times 10^{-6}$  and  $1.2 \times 10^{-6}$ , and the theoretical contribution,  $1.0 \times 10^{-6}$ . This is in good agreement with the four previous measurements of the proton-to-electron mass ratio<sup>21–24</sup> (Fig. 4) and has a similar experimental precision. The most precise value for protons is currently obtained by comparing the  $g$  factors of hydrogen-like  $^{12}\text{C}^{5+}$  and  $^{16}\text{O}^{7+}$  ions measured by the GSI-Mainz collaboration<sup>23,24</sup> with high-field QED calculations. The CODATA recommended value for  $M_p/m_e$  is taken as the average over these experiments. This ratio may be determined to higher precision in the future by laser spectroscopy experiments<sup>25</sup> on  $\text{H}_2^+$  and  $\text{HD}^+$  ions. By assuming<sup>19</sup> CPT invariance, such that  $M_{\bar{p}} = M_p = 1.00727646677(10)$  u, we can further derive the value of  $m_e = 0.0005485799091(7)$  u for the electron mass from our  $\bar{p}\text{He}^+$  result.

The equalities between the antiproton and proton charges and masses, formulated respectively as  $\delta_Q = (Q_p - Q_{\bar{p}})/Q_p$  and  $\delta_M = (M_p - M_{\bar{p}})/M_p$ , have been constrained<sup>26,27</sup> to within  $2 \times 10^{-5}$ . This was achieved by combining X-ray spectroscopic data on antiprotonic atoms ( $\propto Q_p^2 M_{\bar{p}}$ ) with the cyclotron frequency ( $\propto Q_{\bar{p}}/M_{\bar{p}}$ ) of antiprotons confined in Penning traps and measured to a higher precision. We can improve this limit by more than four orders of magnitude by studying the linear dependence<sup>2</sup> of  $\delta_M$  and  $\delta_Q$  on  $v_{\text{th}}$ , that is,  $\delta_M \kappa_M + \delta_Q \kappa_Q \leq |v_{\text{exp}} - v_{\text{th}}|/v_{\text{exp}}$ . For the three transitions, the constants  $\kappa_M$  and  $\kappa_Q$  were estimated<sup>2</sup> to be 2.3–2.8 and 2.7–3.4, respectively. The right-hand side of the inequality was evaluated to be  $<(8 \pm 15) \times 10^{-10}$  by averaging over the three transitions. Furthermore, the constraint that  $(Q_{\bar{p}}/M_{\bar{p}})/(Q_p/M_p) + 1 = 1.6(9) \times 10^{-10}$ , from the TRAP experiment<sup>28,29</sup>, implies that  $\delta_Q \approx \delta_M$ . From this, we conclude that any deviations between the charges and masses are  $<7 \times 10^{-10}$  at the 90% confidence level.

## METHODS SUMMARY

The two continuous-wave seed lasers were stabilized relative to 470-mm-long, monolithic cavities made of ultralow-expansion glass by using the Pound–Drever–Hall technique. The cavities were suspended horizontally by springs and isolated in a vacuum chamber whose temperature was stabilized to  $\pm 0.05^\circ\text{C}$ . Drifts in the laser frequencies were typically  $<0.1 \text{ MHz h}^{-1}$ . The frequency chirp<sup>11,13–15</sup> during pulsed laser amplification was corrected using electro-optic modulators placed inside the pulsed laser resonators, such that its amplitude was reduced to a few megahertz. This remaining chirp was recorded for each laser pulse and its effect corrected for at the data analysis stage. The output beams were

frequency-doubled (second-harmonic generation) or frequency-tripled (third-harmonic generation) to wavelengths of  $\lambda = 264\text{--}417 \text{ nm}$  in  $\beta$ -barium borate and lithium triborate crystals. Simulations<sup>5,13,14</sup> show that additional chirp caused by this frequency conversion is negligible ( $<0.1 \text{ MHz}$ ).

The Cherenkov signals corresponding to  $\bar{p}\text{He}^+$  were recorded using a digital oscilloscope, and the area under the peak in each of these time spectra (Fig. 1b) was plotted as a function of laser frequency to obtain the resonance profiles in Fig. 2. Each data point represents an average of 8–10 antiproton beam arrivals at the target. This measurement was repeated for 10,000 arrivals at various laser intensities, target densities, frequency offsets ( $\nu_d$ ) and alignments of the antiproton beam. The fact that the pulsed laser can maintain absolute precision for the duration of the measurements was verified by using part of the light to measure the 6s–8s two-photon transition frequency of caesium 20 times over a two-week period. The result, with a conservative error of  $1.4 \times 10^{-9}$ , was in good agreement with previous experiments<sup>30</sup>. The acquired resonance profile was fitted with the theoretical two-photon resonance line shape as described in the main text. This a-priori calculation well reproduced the experimental data (Fig. 2). The validity of this method was also partly verified by using it to analyse the above-mentioned caesium two-photon signal<sup>11</sup>.

Received 12 April; accepted 26 May 2011.

- Bluhm, R., Kostecký, V. A. & Russell, N. CPT and Lorentz tests in hydrogen and antihydrogen. *Phys. Rev. Lett.* **82**, 2254–2257 (1999).
- Hayano, R. S., Hori, M., Horváth, D. & Widmann, E. Testing CPT invariance with antiprotonic helium. *Rep. Prog. Phys.* **70**, 1995–2065 (2007).
- Korobov, V. I. Calculations of transitions between metastable states of antiprotonic helium including relativistic and radiative corrections of order  $R_\infty \alpha^4$ . *Phys. Rev. A* **77**, 042506 (2008).
- Hori, M. *et al.* Direct measurement of transition frequencies in isolated  $\bar{p}\text{He}^+$  atoms, and new CPT-violation limits on the antiproton charge and mass. *Phys. Rev. Lett.* **91**, 123401 (2003).
- Hori, M. *et al.* Determination of the antiproton-to-electron mass ratio by precision laser spectroscopy of  $\bar{p}\text{He}^+$ . *Phys. Rev. Lett.* **96**, 243401 (2006).
- Parthey, C. G. *et al.* Precision measurement of the hydrogen-deuterium 1s-2s isotope shift. *Phys. Rev. Lett.* **104**, 233001 (2010).
- Gabrielse, G. *et al.* Antihydrogen production within a Penning-Ioffe trap. *Phys. Rev. Lett.* **100**, 113001 (2008).
- Andresen, G. B. *et al.* Trapped antihydrogen. *Nature* **468**, 673–676 (2010).
- Hori, M. & Korobov, V. I. Calculation of transition probabilities and ac Stark shifts in two-photon laser transitions of antiprotonic helium. *Phys. Rev. A* **81**, 062508 (2010).
- Bjorkholm, J. E. & Liao, P. F. Resonant enhancement of two-photon absorption in sodium vapor. *Phys. Rev. Lett.* **33**, 128–131 (1974).
- Hori, M. & Dax, A. Chirp-corrected, nanosecond Ti:sapphire laser with 6 MHz linewidth for spectroscopy of antiprotonic helium. *Opt. Lett.* **34**, 1273–1275 (2009).
- Udem, Th., Holzwarth, R. & Hänsch, T. W. Optical frequency metrology. *Nature* **416**, 233–237 (2002).
- Meyer, V. *et al.* Measurement of the 1s-2s energy interval in muonium. *Phys. Rev. Lett.* **84**, 1136–1139 (2000).
- Eikema, K. S. E., Ubachs, W., Vassen, W. & Hogervorst, W. Lamb shift measurement in the  $1^1\text{S}$  ground state of helium. *Phys. Rev. A* **55**, 1866–1884 (1997).
- Bakule, P. *et al.* A chirp-compensated, injection-seeded alexandrite laser. *Appl. Phys. B* **71**, 11–17 (2000).
- Korobov, V. I. Hyperfine structure of metastable states in  $^3\text{He}^+\bar{p}$ . *Phys. Rev. A* **73**, 022509 (2006).
- Bakalov, D., Jeziorski, B., Korona, T., Szałewicz, K., Tchoukova, E. Density shift and broadening of transition lines in antiprotonic helium. *Phys. Rev. Lett.* **84**, 235–238 (2000).
- Hori, M. *et al.* Sub-ppm laser spectroscopy of antiprotonic helium and a CPT-violation limit on the antiprotonic charge and mass. *Phys. Rev. Lett.* **87**, 093401 (2001).
- Mohr, P. J. & Taylor, B. N. CODATA recommended values of the fundamental constants: 2002. *Rev. Mod. Phys.* **77**, 1–107 (2005).
- Pohl, R. *et al.* The size of the proton. *Nature* **466**, 213–216 (2010).
- Van Dyck, R. S. Jr, Moore, F. L., Farnham, D. L. & Schwinberg, P. B. Improved measurement of proton-electron mass ratio. *Bull. Am. Phys. Soc.* **31**, 244–244 (1986).
- Farnham, D. L., Van Dyck, R. S. Jr & Schwinberg, P. B. Determination of the electron's atomic mass and the proton/electron mass ratio via Penning trap mass spectroscopy. *Phys. Rev. Lett.* **75**, 3598–3601 (1995).
- Beier, T. *et al.* New determination of the electron's mass. *Phys. Rev. Lett.* **88**, 011603 (2002).
- Verdú, J. *et al.* Electronic  $g$  factor of hydrogenlike oxygen  $^{16}\text{O}^{7+}$ . *Phys. Rev. Lett.* **92**, 093002 (2004).
- Koelmeij, J. C. J., Roth, B., Wicht, A., Ernsting, I. & Schiller, S. Variational spectroscopy of  $\text{HD}^+$  with 2-ppb accuracy. *Phys. Rev. Lett.* **98**, 173002 (2007).
- Hughes, R. J. & Deutch, B. I. Electric charges of positrons and antiprotons. *Phys. Rev. Lett.* **69**, 578–581 (1992).
- Nakamura, K. *et al.* Review of particle physics. *J. Phys. G* **37**, 075021 (2010).
- Gabrielse, G. *et al.* Precision mass spectroscopy of the antiproton and proton using simultaneously trapped particles. *Phys. Rev. Lett.* **82**, 3198–3201 (1999).

29. Thompson, J. K., Rainville, S. & Pritchard, D. E. Cyclotron frequency shifts arising from polarization forces. *Nature* **430**, 58–61 (2004).
30. Fendel, P., Bergeson, S. D., Udem, T. & Hänsch, T. W. Two-photon frequency comb spectroscopy of the 6s–8s transition in cesium. *Opt. Lett.* **32**, 701–703 (2007).

**Acknowledgements** This work was supported by the European Science Foundation (EURYI), Monbukagakusho (grant no. 20002003), the Munich Advanced Photonics cluster of the Deutsche Forschungsgemeinschaft, the Hungarian Research Foundation (K72172) and the Austrian Federal Ministry of Science and Research. We thank the CERN Antiproton Decelerator and Proton Synchrotron operational staff, the CERN cryogenics laboratory, J. Alnis, D. Bakalov, J. Eades, R. Holzwarth, V. I. Korobov, M. Mitani, W. Pirkel and T. Udem.

**Author Contributions** M.H. designed the two-photon experiment. M.H. and A.D. developed the laser systems and carried out the caesium and rubidium measurements. M.H. and D.B. constructed the cryogenic target. M.H. developed the antiproton beam profile monitors, Cherenkov counters, cryogenic optics and data acquisition system. D.B. and M.H. wrote the analysis software. All authors contributed to the beam-time data taking and analysis. M.H. wrote the manuscript and all authors discussed the results and contributed to the editing.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.H. ([masaki.hori@mpq.mpg.de](mailto:masaki.hori@mpq.mpg.de)).



# Carbon loss from an unprecedented Arctic tundra wildfire

Michelle C. Mack<sup>1</sup>, M. Syndonia Bret-Harte<sup>2</sup>, Teresa N. Hollingsworth<sup>3</sup>, Randi R. Jandt<sup>4</sup>, Edward A. G. Schuur<sup>1</sup>, Gaius R. Shaver<sup>5</sup> & David L. Verbyla<sup>6</sup>

Arctic tundra soils store large amounts of carbon (C) in organic soil layers hundreds to thousands of years old that insulate, and in some cases maintain, permafrost soils<sup>1,2</sup>. Fire has been largely absent from most of this biome since the early Holocene epoch<sup>3</sup>, but its frequency and extent are increasing, probably in response to climate warming<sup>4</sup>. The effect of fires on the C balance of tundra landscapes, however, remains largely unknown. The Anaktuvuk River fire in 2007 burned 1,039 square kilometres of Alaska's Arctic slope, making it the largest fire on record for the tundra biome and doubling the cumulative area burned since 1950 (ref. 5). Here we report that tundra ecosystems lost  $2,016 \pm 435 \text{ g C m}^{-2}$  in the fire, an amount two orders of magnitude larger than annual net C exchange in undisturbed tundra<sup>6</sup>. Sixty per cent of this C loss was from soil organic matter, and radiocarbon dating of residual soil layers revealed that the maximum age of soil C lost was 50 years. Scaled to the entire burned area, the fire released approximately 2.1 teragrams of C to the atmosphere, an amount similar in magnitude to the annual net C sink for the entire Arctic tundra biome averaged over the last quarter of the twentieth century<sup>7</sup>. The magnitude of ecosystem C lost by fire, relative to both ecosystem and biome-scale fluxes, demonstrates that a climate-driven increase in tundra fire disturbance may represent a positive feedback, potentially offsetting Arctic greening<sup>8</sup> and influencing the net C balance of the tundra biome.

The Arctic terrestrial C cycle has the potential to influence global climate through feedbacks to recent warming, but the net sign, magnitude and pace of these feedbacks remains uncertain<sup>7,9,10</sup>. Wildfire is a climate-sensitive process that can rapidly transfer large stocks of terrestrial C to the atmosphere, reduce landscape C storage and alter ecosystem C balance<sup>11</sup>. In the boreal forest biome, where fire has been part of the historic disturbance regime, burned area has increased over the past several decades, concurrently with summer warming and drying<sup>12,13</sup>. Arctic tundra has experienced similar climatic changes in the past two decades<sup>14</sup> and although the tundra has little history of fire, models<sup>15,16</sup> predict increasing fire here if the climate continues to warm. Recent observations already indicate increasing burned area in Alaskan Arctic and subarctic tundra regions over the past two decades, even though large fires have been infrequent since the early Holocene<sup>3,4</sup>.

Little is known about the impacts of fire on the carbon stocks and cycling of tundra ecosystems that have no recent history of fire disturbance. Cool, wet soils underlain by permafrost (permanently frozen ground) may restrict burning to plant litter and aboveground biomass, causing relatively little change in ecosystem C pools. However, Arctic tundra landscapes store large pools of C in organic-rich surface soil horizons that have accumulated over millennia<sup>2</sup>, and peatland organic soils, at least in boreal and temperate ecosystems, can be highly flammable when dry and may burn deeply<sup>17,18</sup>. As much as 18% of

the one-metre-depth soil C pool in Arctic tundra may be vulnerable to direct loss through combustion because it is contained in surface organic horizons<sup>1</sup>. This proportion has been intensively quantified for the 100 Pg C stored in North American Arctic soils<sup>1</sup>, and is probably similar for the 496 Pg C stored across the circum-Arctic permafrost region<sup>19</sup>. Thus, if tundra soils burn, increasing fire in this biome could rapidly transfer a substantive amount of C to the atmosphere, reducing landscape C storage, increasing atmospheric CO<sub>2</sub> levels, and amplifying climate warming.

Beyond direct emissions of C, disturbance by fire has the potential to cause threshold changes in ecosystem structure and function by altering the surface reflectance (albedo) and energy balance of landscapes underlain by permafrost<sup>20</sup>. Fire effects on albedo can be variable and, in some cases, are one of the few factors that have been shown to offset positive climate forcing from fire-driven C loss<sup>21</sup>. But fire also removes the soil organic layer that insulates permafrost from warm summer temperatures, leading to thaw and destabilization of the ground surface and presumably altered ecosystem C balance<sup>9</sup>. Processes that destabilize permafrost expose organic C deep in the soil profile to increased decomposition and future release to the atmosphere<sup>22</sup>, further amplifying the positive feedback to high-latitude warming.

In the late summer and autumn of 2007, the Anaktuvuk River fire burned 1,039 km<sup>2</sup> of Arctic tundra on the North Slope of the Brooks Range, Alaska, USA, doubling the cumulative area burned in this region over the past 50 years (Supplementary Fig. 1 and Supplementary Table 1). This fire was an order of magnitude larger than the average fire size in the historic record for the North Slope (Supplementary Table 1). Sediment cores from lakes within the burn perimeter showed no evidence of regional fires in the past 5,000 years (ref. 4). Here we report results from a 2008 field campaign to examine C loss from the vegetation and soils of moist acidic tundra ecosystems that covered 54% of the area burned in the fire. We examined concurrent nitrogen (N) loss because this is the element most likely to limit plant productivity in this system and thus its loss could constrain post-fire C uptake over longer timescales. We developed a biometric method for estimating pre-fire soil organic layer depth in burned tundra, based on the persistent crowns of the dominant, tussock-forming sedge species, *Eriophorum vaginatum* L., which survived to re-sprout after fire (Supplementary Methods). We measured post-fire residual soil organic matter and reconstructed fire-driven losses of C and N in 20 sites widely distributed across the burn (Supplementary Fig. 3). The radiocarbon age of the residual soil surface was measured to determine whether the fire burned into soil C pools that were hundreds or even thousands of years old (Supplementary Methods).

Reconstructed pre-fire soil organic layer C pools averaged 7,682 g C m<sup>-2</sup> (Table 1) across moist acidic tundra in the Anaktuvuk River fire scar, a value similar to the average (7,500 g C m<sup>-2</sup>) reported for upland tundra sites throughout the North American Arctic<sup>1</sup>. Fire

<sup>1</sup>Department of Biology, University of Florida, PO Box 118525, Gainesville, Florida 32611, USA. <sup>2</sup>Institute of Arctic Biology, University of Alaska Fairbanks, PO Box 757000, Fairbanks, Alaska 99775, USA. <sup>3</sup>Boreal Ecology Cooperative Research Unit, PNW Research Station USDA Forest Service, PO Box 756780, University of Alaska Fairbanks, Fairbanks, Alaska 99775, USA. <sup>4</sup>Alaska Fire Service, Bureau of Land Management, PO Box 35005, Fort Wainwright, Alaska 99703, USA. <sup>5</sup>The Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA. <sup>6</sup>Department of Forest Sciences, University of Alaska Fairbanks, PO Box 757200, Fairbanks, Alaska, USA.

**Table 1 | Ecosystem effects of the 2007 Anaktuvuk River fire**

	Statistics	Soil depth (cm)	C pool (g C m <sup>-2</sup> )			N pool (g N m <sup>-2</sup> )		
			Soil	Plant	Total	Soil	Plant	Total
Pre-fire (measured)	Mean (s.e.m.)	21.5 (1.5)	7,682 (766)	559 (33)	8,241 (764)	287 (36)	17 (1)	303 (35)
	Min	12.3	2,842	355	3,197	84	9	93
	Max	43.3	19,470	840	19,833	847	28	856
Post-fire (reconstructed)	Mean (s.e.m.)	15.4 (1.1)	6,354 (613)	0	6,354 (613)	245 (34)	0	245 (34)
	Min	8.5	2,718	0	2,718	67	0	67
	Max	28.0	12,007	0	12,007	624	0	624
Loss (reconstructed)	Mean (s.e.m.)	6.1 (1.0)	1,456 (445)	559 (33)	2,016 (435)	47 (18)	17 (1)	64 (17)
	Min	3.4	379	355	753	9	9	19
	Max	22.8	9,460	840	9,823	370	28	379

Measured or reconstructed soil organic layer depth, and soil and combustible aboveground plant C and N pools in 20 upland moist acidic tundra sites that burned in the 2007 Anaktuvuk River fire, North Slope of the Brooks Range, Alaska, USA.

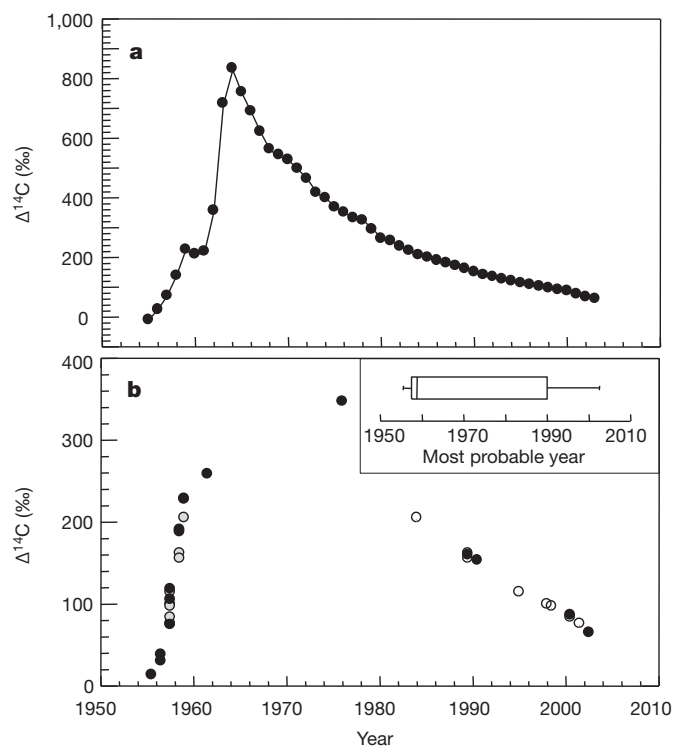
reduced combustible ecosystem C pools by  $31 \pm 3\%$  (mean  $\pm 1$  standard error) resulting in a net loss of about  $2,016 \text{ g C m}^{-2}$  and  $64 \text{ g N m}^{-2}$  (Table 1). This is in the mid-range of emissions reported for North American boreal forest fires<sup>15,17,18</sup>. On average,  $62 \pm 4\%$  of ecosystem C and a similar percentage of N loss came from soil, with the remainder from plant biomass. The depth of the soil organic layer was reduced by  $30 \pm 3\%$  and residual organic layer depth, a factor likely to influence insulation of permafrost, averaged 15 cm (Table 1). Survivorship of *E. vaginatum* tussocks averaged  $87 \pm 11\%$ , while the mosses, lichens and dwarf shrubs that inhabited spaces among tussocks before the fire were almost completely combusted.

Despite wide variation in pre-fire soil organic layer depth (12.3–43.3 cm) and organic layer carbon pools ( $2,842\text{--}19,470 \text{ g C m}^{-2}$ ), the depth burned and C and N lost were relatively constant across sites. Pre-fire soil organic layer depth was not related to depth lost but was highly predictive of post-fire residual depth ( $R^2 = 0.82$ ,  $P < 0.001$ ,  $n = 20$  sites). This is consistent with observations of organic soil burning in Alaskan boreal forest<sup>17</sup> and suggests that in these structured organic soils, burning is restricted to surface soil by a mechanism independent of potential fuel, such as variation with depth of bulk density, particle size and hydraulic connectivity with the water table. Our results indicate that residual soil organic layer depth is not an adequate proxy for fire severity, depth of burning, or C emissions from tundra soils because of the large variation in pre-fire depth. As a result, estimates of emissions that index consumption as a fraction of pre-fire C pool size will overestimate consumption in deep organic soils.

Age and density of soil C typically increase with depth in peat soils<sup>23</sup>. Plant production continuously adds new material at the surface of the soil while decomposition and leaching remove organic matter throughout the soil profile, resulting in older, deeper strata that are thinner and higher in density than younger strata near the surface. Radiocarbon dating of residual soil surfaces showed that burning in the Anaktuvuk River fire was limited to younger, upper soil layers. All samples of moss macrofossils removed from the surface soils of the burned sites contained bomb-enriched radiocarbon, setting the maximum age of lost soil C at 50 years (Fig. 1). Across the 20 soil profiles sampled for radiocarbon, the fire burned down to strata with an average year of formation of 1970 (95% confidence intervals: 1959–1981). The relationship between age and depth—6.1 cm and 37 years—observed here is similar to the accumulation rates of surface organic soils elsewhere on the North Slope<sup>24</sup> and Canada<sup>23</sup>. The profile that had the deepest burning (15 cm lost) also had the oldest burned surface date (1954), but there was no evidence that the fire burned into old, pre-bomb layers in any of the profiles despite the existence of organic soil layers  $>30$  cm deep.

On a per unit area burned basis, C and N losses from the Anaktuvuk River fire were substantially larger than annual input or output fluxes of undisturbed tundra. Carbon loss was approximately five times

greater than annual net primary productivity of moist acidic tundra ( $450 \text{ g C m}^{-2}$ ; ref. 25) and two orders of magnitude larger than the annual balance between production and decomposition (net ecosystem carbon exchange,  $29.5 \text{ g C m}^{-2} \text{ yr}^{-1}$ ; ref. 6). Nitrogen deposition rates in Arctic Alaska are low ( $0.008\text{--}0.06 \text{ g N m}^{-2} \text{ yr}^{-1}$ ; ref. 26) and inputs from biological N fixation are estimated to be slightly higher: at most,  $0.1 \text{ g N m}^{-2} \text{ yr}^{-1}$  in unburned tundra<sup>26</sup>. Assuming steady-state dynamics, the fire emitted about 400 years of N accumulation.



**Figure 1 | Atmospheric radiocarbon values over the past 57 years and radiocarbon values of the burned soil surface in the Anaktuvuk River fire scar, Alaska, USA.** **a**, Atmospheric radiocarbon content versus time (adapted from ref. 47 in the Supplementary Methods). **b**, Radiocarbon content of moss macrofossils removed from the burned surface versus estimated age of  $n = 20$  profiles sampled in the Anaktuvuk River fire scar. Black circles indicate samples where location on the bomb curve (ascending or descending) was confirmed by measuring two depths (0–1 cm and 1–2 cm), while grey or white circle pairs indicate the two possible ages of samples where one depth (0–1 cm) was measured (Supplementary Methods). The inset shows a box plot of median most probable age (vertical line in rectangle), the range within which the central 50% of the values fall (length of rectangle), hinges (first and third quartiles) and whiskers (range of values within 1.5 Hs of the hinges).

Scaled to the 1,039 km<sup>2</sup> burned area, the fire released approximately 2.1 Tg C to the atmosphere. Extrapolating from moist acidic tundra alone excludes the potential for different emissions from shrub tundra, which occupied 30% of the vegetated burn area. Higher aboveground biomass in shrubs, however, may have been offset by the presence of fire-resistant woody stems and lower density and shallower depth of organic soil<sup>27</sup>, such that emissions per unit area may not be widely different. Carbon emissions from the entire burn were then similar in magnitude to the annual net C sink simulated by process models for the entire Arctic tundra biome (3–4 Tg C yr<sup>-1</sup>) over the last quarter of the twentieth century<sup>7</sup>.

Our results show how rapidly a single tundra fire—even one that burned relatively superficially—can offset local and biome-scale C uptake. The rate of soil C loss per unit burned area was more rapid than other climate-driven pathways of loss that may affect tundra ecosystem C balance in the future. Combustion losses were 30–50 times more rapid than decomposition-related mechanisms such as the stimulation of soil organic matter decomposition from a 5 °C increase in mean annual temperature<sup>28</sup>, temperature-driven increases in N availability that can directly stimulate microbial activity<sup>29</sup>, or ground subsidence as a consequence of permafrost thaw and melting ground ice<sup>22</sup>. At both local and regional scales, fire emissions of this magnitude can instantaneously offset or reverse C cycling processes hypothesized to feedback negatively to warming such as greening of the Arctic<sup>8,30</sup>.

A novel disturbance of this magnitude also has implications for future interactions between ecosystems and climate change. The relatively young age of soil C lost and the high survivorship of graminoids (sedges) in the Anaktuvuk River fire create the potential for plant and soil C losses to be replaced by new plant production before the next fire, assuming a fire return interval of 80–150 years, similar to that of boreal forest south of the Brooks Range<sup>12,15</sup>. This replacement time implies, however, that transition to a new fire regime would catalyse a net transfer of C from ecosystems to the atmosphere as subsequent fires shift never-burned tundra to early-successional ecosystems with lower C storage.

Another future scenario is that fire will act as a mechanism for threshold change such that graminoid tundra is replaced by shrub tundra after fire. A shift to shrub tundra has been independently occurring across Arctic Alaska<sup>30</sup>, and vegetation transition from graminoid to shrub domination has been linked to fire in some palaeoecological studies<sup>3</sup>. This vegetation change could trigger additional positive feedbacks to climate warming because shrub-dominated ecosystems have higher productivity and plant biomass<sup>25</sup> offset by lower soil C stocks<sup>27</sup>, in addition to direct regional warming induced by the lower albedo of shrubs than of graminoid tundra<sup>14</sup>.

Ecosystems also interact directly with climate through surface reflectance and energy partitioning. Alteration of the tundra surface by the Anaktuvuk River fire decreased albedo by 50–71% and increased partitioning of net radiation into ground heat flux<sup>31</sup>. Fire-induced changes in ecosystem albedo were, in the case of a boreal forest, found to offset other positive climate forcing factors over 80 years of succession, including changes in ecosystem C storage and initial post-fire decreases in albedo<sup>21</sup>. This negative feedback from albedo forcing, however, was largely driven by changes in tree canopy shading of spring snow, a factor unlikely to be important in moist acidic tundra ecosystems where few plant species extend above the snow pack. A more likely possibility for tundra is that the increased ground heat flux initiated by decreased albedo and reduced organic soil insulation will destabilize permafrost, leading to internal feedbacks that cause ground ice to melt and surfaces to subside<sup>9</sup>. Over centennial timescales, deep permafrost C pools have the potential for C loss an order of magnitude larger than the C gained in plant biomass by the transition to shrub tundra or even boreal forest<sup>19,22</sup>.

The half-century record of Arctic and subarctic tundra area burned in Alaska points to increasing fire in tundra regions over the past two

decades, concurrent with increasing mean annual temperature and sea ice retreat<sup>4</sup>, trends that seem likely to continue<sup>10</sup>. Our observations of C loss from the Anaktuvuk River fire support the idea that tundra fires have the potential to release large amounts of C and decrease landscape C stocks, having an immediate impact on atmospheric C and climate that is rapid in comparison to other potential climate change feedbacks from tundra ecosystems. Although the future trajectory of ecosystem recovery in the Anaktuvuk River fire scar is yet unknown, most of the plausible short- and long-term scenarios lead to amplification of climate warming.

## METHODS SUMMARY

The Anaktuvuk River fire scar is located on the North Slope of the Brooks Range, Alaska, approximately 23 km northwest of Toolik Field Station (68.5833 °N, 149.7167 °W). This region is underlain by permafrost. The mean annual temperature is approximately –10 °C and mean annual precipitation is 30 cm. Before the fire, 54% of the vegetated area within the burn perimeter was classified as upland moist acidic tundra (MAT; soil pH < 5.5), 15% as moist non-acidic tundra (soil pH > 5.5), and 30% as shrubland. MAT is pan-Arctic in distribution and covers as much as 336 × 10<sup>6</sup> km<sup>2</sup> of the tundra biome<sup>32</sup>. We focused our study on MAT because of its widespread distribution and because crowns of the dominant plant species, *Eriophorum vaginatum*, persisted through fire and provided a benchmark of pre-fire soil organic matter depth and plant biomass. We sampled twenty MAT sites within the burn in July and August of 2008. Eleven widely distributed unburned sites were sampled to develop and test regional biometric relationships between *E. vaginatum* meristem height above the mineral soil versus pre-fire soil organic layer depth, and depth versus bulk density, [C] or [N]. We modified a boreal forest method<sup>17</sup> for use in MAT and used it to calculate site-specific estimates of pre- and post-fire soil organic matter pools and estimate losses. We used the bomb peak method to radiocarbon date the post-fire soil surface and examine whether the fire burned into old and probably irreplaceable soil C pools. Further description of the region, sites and methods used are included in the Supplementary Information.

Received 11 November 2010; accepted 10 June 2011.

- Ping, C. L. *et al.* High stocks of soil organic carbon in the North American Arctic region. *Nature Geosci.* **1**, 615–619 (2008).
- Harden, J. W., Sundquist, E. T., Stallard, R. F. & Mark, R. K. Dynamics of soil carbon during deglaciation of the Laurentide ice-sheet. *Science* **258**, 1921–1924 (1992).
- Higuera, P., Brubaker, L. B., Anderson, P. M., Brown, T. A. & Kennedy, A. T. Frequent fires in ancient shrub tundra: implications of paleorecords for Arctic environmental change. *PLoS ONE* **3**, e0001744 (2008).
- Hu, F. S. *et al.* Tundra burning in Alaska: linkages to climatic change and sea ice retreat. *J. Geophys. Res. Biogeosci.* **115**, G04002 (2010).
- Jones, B. M. *et al.* Fire behavior, weather, and burn severity of the 2007 Anaktuvuk river tundra fire, North Slope, Alaska. *Arct. Antarct. Alp. Res.* **41**, 309–316 (2009).
- Oechel, W. C. *et al.* A scaling approach for quantifying the net CO<sub>2</sub> flux of the Kuparuk river basin, Alaska. *Glob. Change Biol.* **6**, 160–173 (2000).
- McGuire, A. D. *et al.* Sensitivity of the carbon cycle in the Arctic to climate change. *Ecol. Monogr.* **79**, 523–555 (2009).
- Goetz, S. J., Bunn, A. G., Fiske, G. J. & Houghton, R. A. Satellite-observed photosynthetic trends across boreal North America associated with climate and fire disturbance. *Proc. Natl Acad. Sci. USA* **102**, 13521–13525 (2005).
- Schuur, E. A. G. *et al.* Vulnerability of permafrost carbon to climate change: Implications for the global carbon cycle. *Bioscience* **58**, 701–714 (2008).
- The Intergovernmental Panel on Climate Change (IPCC) *Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC* (Cambridge Univ. Press, 2007).
- Field, C. B., Lobell, D. B., Peters, H. A. & Chiariello, N. R. Feedbacks of terrestrial ecosystems to climate change. *Annu. Rev. Environ. Resour.* **32**, 1–29 (2007).
- Kasischke, E. S. & Turetsky, M. R. Recent changes in the fire regime across the North American boreal region—spatial and temporal patterns of burning across Canada and Alaska. *Geophys. Res. Lett.* **33** (13) L09703 (2006).
- Zimov, S. A. *et al.* Contribution of disturbance to increasing seasonal amplitude of atmospheric CO<sub>2</sub>. *Science* **284**, 1973–1976 (1999).
- Chapin, F. S. III *et al.* Role of land-surface changes in Arctic summer warming. *Science* **310**, 657–660 (2005).
- Balshi, M. S. *et al.* Vulnerability of carbon storage in North American boreal forests to wildfires during the 21st century. *Glob. Change Biol.* **15**, 1491–1510 (2009).
- Krawchuck, M. A., Moritz, M. A., Parisien, M.-A., Van Dorn, J. & Hayhoe, K. Global pyrogeography: the current and future distribution of wildfire. *PLoS ONE* **4**, 1–12 (2009).
- Boby, L. A., Schuur, E. A. G., Mack, M. C., Johnstone, J. F. & Verbyla, D. L. Quantifying fire severity, carbon and nitrogen emissions in Alaska's boreal forests. *Ecol. Appl.* **20**, 1633–1647 (2010).
- Turetsky, M. R. *et al.* Recent acceleration of biomass burning and carbon losses in Alaskan forests and peatlands. *Nature Geosci.* **4**, 27–31 (2011).



19. Tarnocai, C. *et al.* Soil organic carbon pools in the northern circumpolar permafrost region. *Glob. Biogeochem. Cycles* **23**, GB2023 1–11 (2009).
20. Chapin, F. S. III *et al.* Arctic and boreal ecosystems of western North America as components of the climate system. *Glob. Change Biol.* **6**, 211–223 (2000).
21. Randerson, J. T. *et al.* The impact of boreal forest fire on climate warming. *Science* **314**, 1130–1132 (2006).
22. Schuur, E. A. G. *et al.* The effect of permafrost thaw on old carbon release and net carbon exchange from tundra. *Nature* **459**, 556–559 (2009).
23. Trumbore, S. E. & Harden, J. W. Accumulation and turnover of carbon in organic and mineral soils of the BOREAS northern study area. *J. Geophys. Res.* **102**, 28817–28830 (1997).
24. Schell, D. & Barnett, B. *Peat Cores from the Toolik Lake and Imnaviat Creek Region* (<http://metacat.lternet.edu/das/lter/advancedsearch.jsp?site=ARC>) (12 files named 89spt01.txt to 89spt12.txt) (US Long Term Ecological Research Database, 1989).
25. Shaver, G. R. & Chapin, F. S. III. Production:biomass relationships and element cycling in contrasting Arctic vegetation types. *Ecol. Monogr.* **61**, 1–31 (1991).
26. Hobara, S. *et al.* Nitrogen fixation in surface soils and vegetation in an Arctic tundra watershed: a key source of atmospheric nitrogen. *Arct. Antarct. Alp. Res.* **38**, 363–372 (2006).
27. Ping, C. L., Bockheim, J. G., Kimble, J. M., Michaelson, G. J. & Walker, D. A. Characteristics of cryogenic soils along a latitudinal transect in Arctic Alaska. *J. Geophys. Res. Atmos.* **103** (D22), 28917–28928 (1998).
28. Luo, Y. Q. Terrestrial carbon-cycle feedback to climate warming. *Annu. Rev. Ecol. Syst.* **38**, 683–712 (2007).
29. Mack, M. C., Schuur, E. A. G., Bret-Harte, M. S., Shaver, G. R. & Chapin, F. S. III. Ecosystem carbon storage in Arctic tundra reduced by long-term nutrient fertilization. *Nature* **431**, 440–443 (2004).
30. Sturm, M., Racine, C. & Tape, K. Climate change—increasing shrub abundance in the Arctic. *Nature* **411**, 546–547 (2001).
31. Rocha, A. V. & Shaver, G. R. Advantages of a two band EVI calculated from solar and photosynthetically active radiation fluxes. *Agric. For. Meteorol.* **149**, 1560–1563 (2009).
32. Walker, D. A. *et al.* The circumpolar Arctic vegetation map. *J. Veg. Sci.* **16**, 267–282 (2005).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank J. Ahgook Jr, L. Boby, M. Cahill, E. Miya, E. Miller, J. Oyler, C. Roberts, E. Sironen, C. Wachs, C. Wasykowski and D. Yokel for their contributions to fieldwork, C. Apodaca, G. Blohm, E. Brown, G. Crummer and D. Nossov for their contributions to laboratory work and sample analyses, H. Alexander for contributing to data analyses, and P. Ray for insights into tussock morphology. This research was supported by the US NSF Division of Environmental Biology, the Division of Biological Infrastructure and Office of Polar Programs, by the US National Center for Ecological Analysis and Synthesis and by the US Bureau of Land Management Alaska Fire Service and Arctic Field Office.

**Author Contributions** M.C.M., M.S.B.-H., T.N.H., R.R.J. and D.L.V. designed the study with input from E.A.G.S. and G.R.S. M.C.M., T.N.H., R.R.J. and M.S.B.-H. conducted soil and vegetation sampling fieldwork and M.C.M., E.A.G.S. and D.L.V. analysed samples and data. M.C.M. wrote the manuscript with input from all co-authors.

**Author Information** The data described in this study is publicly available in the Arctic Long Term Ecological Research data archive (<http://ecosystems.mbl.edu/arc/burn/data.html>). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.C.M. ([mcmack@ufl.edu](mailto:mcmack@ufl.edu)).

# Transforming binding affinities from three dimensions to two with application to cadherin clustering

Yinghao Wu<sup>1,2,3</sup>, Jeremie Vendome<sup>1,2,3</sup>, Lawrence Shapiro<sup>1,4</sup>, Avinoam Ben-Shaul<sup>5</sup> & Barry Honig<sup>1,2,3</sup>

Membrane-bound receptors often form large assemblies resulting from binding to soluble ligands, cell-surface molecules on other cells and extracellular matrix proteins<sup>1</sup>. For example, the association of membrane proteins with proteins on different cells (*trans*-interactions) can drive the oligomerization of proteins on the same cell<sup>2</sup> (*cis*-interactions). A central problem in understanding the molecular basis of such phenomena is that equilibrium constants are generally measured in three-dimensional solution and are thus difficult to relate to the two-dimensional environment of a membrane surface. Here we present a theoretical treatment that converts three-dimensional affinities to two dimensions, accounting directly for the structure and dynamics of the membrane-bound molecules. Using a multiscale simulation approach, we apply the theory to explain the formation of ordered, junction-like clusters by classical cadherin adhesion proteins. The approach features atomic-scale molecular dynamics simulations to determine interdomain flexibility, Monte Carlo simulations of multidomain motion and lattice simulations of junction formation<sup>3</sup>. A finding of general relevance is that changes in interdomain motion on *trans*-binding have a crucial role in driving the lateral, *cis*-, clustering of adhesion receptors.

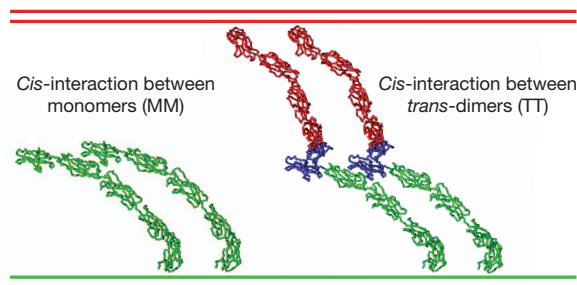
It is commonplace to characterize binding between macromolecules quantitatively by measuring dissociation constants in solution,  $K_d^{(3D)}$ , which are typically defined in three-dimensional (3D) concentration units (for example moles per litre). However, phenomena that take place on membrane surfaces are dependent on two-dimensional (2D) densities and the relevant dissociation constants,  $K_d^{(2D)}$ , are defined in units such as molecules per square micrometre. Measurements of  $K_d^{(2D)}$  are difficult to perform and have only been carried out in a small number of cases<sup>4,5</sup>. Thus, it would be extremely valuable to have a method that could transform measured values of  $K_d^{(3D)}$  into corresponding values of  $K_d^{(2D)}$ . A reasonable simplifying assumption in such a method is that the binding interface formed by any two molecules is essentially identical in 3D and in 2D. The difference in the dissociation constants then results only from the change in dimensionality and from any other effects that arise from the constrained environment of a planar system.

It is possible to transform between two and three dimensions through the simple expression  $K_d^{(2D)} = hK_d^{(3D)}$ , where  $h$  is the 'confinement length'<sup>6,7</sup>. The basic idea is that if two interacting species are confined to a region of length  $h$  along an axis perpendicular to the plane of a membrane, then they are effectively confined to a volume  $Ah$ , where  $A$  is the area per molecule<sup>6–8</sup>. This simple procedure turns a 2D system into a 'quasi-3D system' because there is now a volume associated with each molecule even when it is constrained to a planar membrane. The extent of motion in the third dimension can arise from different factors such as molecular flexibility, rotations with respect to the membrane plane, membrane fluctuations and translational motion of the membranes themselves. A number of studies have used measured 3D and 2D affinities to determine  $h$  for individual systems. However, widely discrepant values have been obtained from the use of different methods to measure 2D affinities<sup>5</sup>; for example, fluorescence measurements

typically yield values for  $h$  of the order of nanometres, whereas mechanical measurements have yielded values for  $h$  of the order of micrometres<sup>5</sup>. Here we focus on cases where two flat, parallel membranes are separated by a distance that allows proteins located on opposing surfaces to interact in *trans* and where proteins located on the same surface oligomerize in *cis*. The values of  $h$  that we find are of the order of nanometres, as is consistent with fluorescence measurements of 2D affinities<sup>5</sup>.

Our specific focus is on the formation of ordered structures by the type I family of classical cadherins. Cadherins have five extracellular immunoglobulin-fold (EC) domains but the *trans*-binding interface is localized entirely on the membrane-distal EC1 domain<sup>9</sup>. We have recently shown that a molecular layer seen in crystal structures of classical cadherins corresponds to the extracellular structure of adherens junctions<sup>10</sup>. In addition to the *trans*-interface, a second, *cis*-, interface is formed between the EC1 domain of one cadherin and a region comprising parts of the EC2 and EC3 domains of another (Fig. 1). Cadherin *trans*-binding affinities have been measured in 3D solution<sup>11</sup>; binding affinities of *cis*-interactions are too weak to measure but have an upper limit of about 1 mM (ref. 10). We use this well-defined system as a basis for the development of general theoretical expressions that accomplish the transformation from 3D to 2D. These expressions, when used in conjunction with experimental data and our multiscale simulations, provide a detailed description of the structural and energetic basis of junction formation and elucidate new principles that are likely to be relevant to other systems.

Figure 2a describes the *trans*-dimerization reaction when cadherins are restricted to the membrane surface. As mentioned above, we assume that the binding interfaces are the same in solution and on a membrane surface, such that the energetic contributions to binding are identical:  $\Delta E(3D) = \Delta E(2D)$ . Hence, the difference in the binding affinities is entirely entropic. Because the *trans*-dimerization interface is located

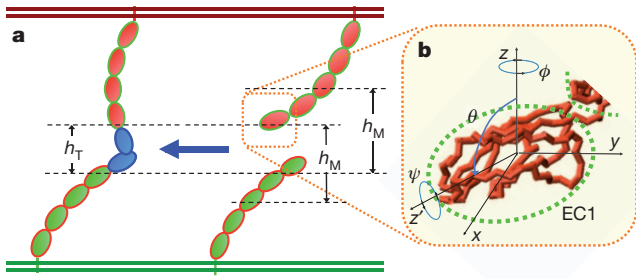


**Figure 1 | Structures of *cis*-dimers formed from cadherin monomers and from *trans*-dimers.** All coordinates are taken from the crystal structure of C-cadherin ectodomains<sup>15</sup>. *Trans*-dimers are formed by EC1 domains (blue) of cadherin monomers from lower (green) and upper (red) cell surfaces. Note that each *trans*-dimer structure has only a single *cis*-interface because the binding regions of the two monomers in a *trans*-dimer face in different directions. This property allows the formation of a 2D lattice in which each pair of *trans*-dimers makes only a single *cis*-interaction<sup>3,10,15</sup>.

<sup>1</sup>Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA. <sup>2</sup>Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA.

<sup>3</sup>Center for Computational Biology and Bioinformatics, Columbia University, New York, New York, 10032, USA. <sup>4</sup>Edward S. Harkness Eye Institute, Columbia University, New York, New York 10032, USA.

<sup>5</sup>Institute of Chemistry and the Fritz Haber Research Center, The Hebrew University, Jerusalem 91904, Israel.



**Figure 2 | Essential coordinates that characterize the dimerization processes of classical cadherins in a 2D membrane environment.** **a**, The five domains of cadherin's extracellular regions are represented by ellipsoids. *Trans*-dimers (blue) can be formed from two cadherin monomers from two apposing cell surfaces. The molecules are free to diffuse in only two dimensions and rotational motion is constrained. A third dimension is introduced through variations in the perpendicular displacement from the membrane surface, defined by the variable  $h$ , which is different for the monomer and the *trans*-dimer. In general,  $h_M$  will be larger than  $h_T$  because *trans*-binding will limit molecular motion. **b**, The rotational degrees of freedom for EC1 domains are characterized by the three Euler angles,  $\phi$ ,  $\theta$  and  $\psi$ .

on EC1, the difference between 3D and 2D affinities is related to the probability that two EC1 domains will encounter one another in an orientation that allows binding. This in turn depends on the local concentration of EC1 domains and on their freedom of rotational motion. As indicated in the figure, we use  $h_M$  and  $h_T$  to denote the ranges of EC1 motion normal to the membrane plane corresponding to monomeric and *trans*-dimeric cadherins, respectively. Thus, unlike in the expression  $K_d^{(2D)} = hK_d^{(3D)}$  (refs 6, 7), we allow for different values of the monomer–dimer confinement length and, hence, their local concentrations, a factor that will prove crucial in the discussion below. To calculate  $h_M$  and  $h_T$ , we make the simplifying assumption that the two adhering membranes are flat and parallel to each other, as illustrated in Fig. 2. Assuming a cadherin density of 80 molecules per square micrometre<sup>11</sup>, the lateral intermolecular distance is about 100 nm (and becomes much smaller once clustering begins). Estimates based on bending rigidity suggest that, over this lateral distance range, spontaneous fluctuations in membrane height are typically only a fraction of a nanometre<sup>12,13</sup>, which is significantly less than the variations in  $h$  due to molecular flexibility considered in this work. Cells *in vivo* can extend membranous protrusions such as filopodia, which on some scale are not flat. Consideration of such issues goes beyond the scope of the current work; however, the treatment given here should provide a good starting point for these more complex instances.

The factors that enter into our treatment of rotational motion are shown in Fig. 2b, where the orientation of the EC1 binding site is described in terms of the three Euler angles,  $\phi$ ,  $\theta$  and  $\psi$ . In 3D, all three rotational angles are unrestricted. By contrast, there are restrictions on the rotational freedom of the membrane-bound molecules except for rotations in  $\phi$ , which corresponds to motion around the  $z$  axis. The rotational entropy is related to the integral over the three Euler angles<sup>14</sup>, which yields  $8\pi^2$  in 3D and a smaller value,  $\Omega < 8\pi^2$ , for membrane-bound molecules (Supplementary Information). Here  $\Omega = (\Delta\omega_M)^2 / \Delta\omega_T$ , where  $\Delta\omega_M = 2\pi\Delta\psi_M[1 - \cos(\Delta\theta_M)]$  and  $\Delta\omega_T = 2\pi\Delta\psi_T[1 - \cos(\Delta\theta_T)]$  are the rotational phase space volumes of the monomer and the *trans*-dimer, respectively (Supplementary Information). Along with the confinement lengths  $h_M$  and  $h_T$ ,  $\Delta\omega_M$  and  $\Delta\omega_T$  describe the ‘confinement’ in rotational motion in the constrained environment of the membrane.

In Supplementary Information, we derive the expression

$$\frac{K_d^{(2D)}(trans)}{K_d^{(3D)}(trans)} = \frac{\Omega}{8\pi^2} \frac{h_M^2}{h_T} = \frac{1}{8\pi^2} \frac{(\Delta\omega_M h_M)^2}{\Delta\omega_T h_T} \quad (1)$$

Equation (1) is quite general, although, as presented here, the variables refer specifically to the EC1 domains of cadherins. We note that it is straightforward to transform from 3D to 2D if  $h_M$ ,  $h_T$ ,  $\Delta\theta_M$ ,  $\Delta\psi_M$ ,  $\Delta\theta_T$  and  $\Delta\psi_T$  are known. These geometric variables will depend on the structures and flexibility of the proteins involved and on the constraints imposed by the membrane environment.

It is instructive to consider the special, hypothetical, case where the reactive EC1 domains of monomers and dimers can freely diffuse within the same ‘reaction’ volume, such that  $h_M = h_T = h$  and, in addition, monomer and dimer rotations in 2D are totally unrestricted, as in 3D ( $\Omega/8\pi^2 = 1$ ). In this case, equation (1) reduces to the simple expression of refs 6, 7, which, however, does not account for real differences in binding free energies in 2D and 3D. Real differences are due to two effects. First, because  $h_M > h_T$  and  $\Delta\omega_M > \Delta\omega_T$ , the volume available to monomers in 2D is larger than that available to *trans*-dimers, implying that the binding affinity is smaller than in the 3D case. Second, the rotational entropy loss on binding in 2D is smaller than that in 3D, as quantitatively represented by  $\Omega/8\pi^2 < 1$ , resulting in enhancement of the binding affinity in 2D relative to 3D. These two effects will thus partly compensate each other, as demonstrated below in quantitative terms based on molecular-level simulations.

As mentioned above, many membrane receptors form lateral clusters on the cell surface driven by the formation of a distinct interprotein *cis*-interface<sup>2</sup>, which for the specific case of cadherins has been characterized crystallographically<sup>10,15</sup>. Asymmetric *cis*-interfaces can form between two monomers, as well as between two *trans*-dimers, as shown in Fig. 1. In Supplementary Information, we derive equations for the 2D dissociation constants appropriate to the *cis*-dimerization of cadherin monomers,  $K_d^{(2D)}{}_{MM}(cis)$ , and *trans*-dimers,  $K_d^{(2D)}{}_{TT}(cis)$ . We show there that

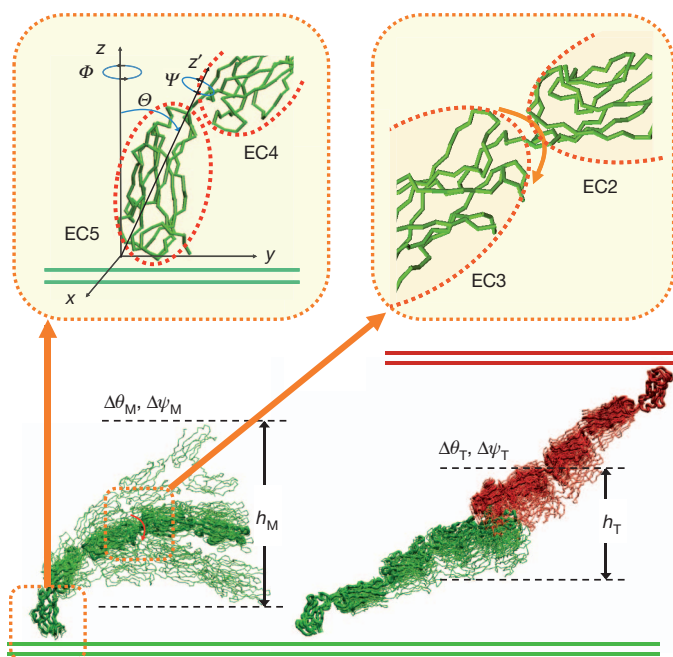
$$\frac{K_d^{(2D)}{}_{MM}(cis)}{K_d^{(2D)}{}_{TT}(cis)} = \left( \frac{\Delta\omega_M h_M}{\Delta\omega_T h_T} \right)^2 \quad (2)$$

Equation (2), which accounts for differences in the strength of *cis*-interactions between monomers and *trans*-dimers, provides physical insights as to the coupling between *trans*- and *cis*-interactions. Even if *cis*-dimers formed from *trans*-dimers have an identical interface to that formed between monomers, the affinities will be different owing to differences in their respective rotational and vibrational flexibilities, as reflected by the factors  $\Delta\omega_M/\Delta\omega_T$  and  $h_M/h_T$ , respectively. Qualitatively, because both factors are larger than one, it follows that the lateral attraction between *trans*-dimers is stronger than that between monomers.

In Methods, we describe a multiscale simulation approach that yields estimates of the six variables,  $h_M$ ,  $h_T$ ,  $\Delta\theta_M$ ,  $\Delta\psi_M$ ,  $\Delta\theta_T$  and  $\Delta\psi_T$ , that define the transformation between 3D and 2D. It is evident from the simulations (Fig. 3) that *trans*- and/or *cis*-dimer formation places significant constraints on the molecular system. Values of  $h$ ,  $\Delta\theta$  and  $\Delta\psi$  are reduced by a factor of approximately two to three in going from a monomer to a *trans*- or *cis*-dimer (that is,  $h_T < h_M$ ,  $\Delta\theta_T < \Delta\theta_M$  and  $\Delta\psi_T < \Delta\psi_M$ ), an effect that will tend to weaken binding affinities (Supplementary Table 1). Supplementary Table 1 also shows 3D and 2D dissociation constants for the dimerization reactions occurring in solution and on the membrane. Notably, the values of  $K_d^{(2D)}$  for *trans*-interactions reported in Supplementary Table 1 (ranging from 15 to 250  $\mu\text{m}^{-2}$ ) for N-cadherin are in the range obtained from measurements on molecules associated with the T-cell system<sup>4,5,16</sup>, whereas those for E-cadherin are about an order of magnitude weaker owing largely to the greater values of  $K_d^{(3D)}$ .

The most drastic effect seen in the simulations is the difference in  $K_d^{(2D)}$  of three to five orders of magnitude for lateral, *cis*-, dimerization affinities between monomers and *trans*-dimers. The increased binding affinity for *trans*-dimers has a clear physical explanation. The association of two cadherin monomers into a *cis*-dimer places severe constraints on the interdomain mobility of both ectodomains, such that





**Figure 3 | Monte Carlo simulations of the flexibility of the cadherin ectodomain.** The rotations of the EC5 domain with respect to the membrane plane depend on the three Euler angles,  $\Phi$ ,  $\Theta$  and  $\Psi$ , of that domain, as shown in the upper left panel. The interdomain hinge motion, indicated by a red arrow, is shown in the upper right panel. The lower part of the figure shows the superposition of different monomer and *trans*-dimer conformations generated by the simulations. The range of values for  $h$ ,  $\Delta\psi$  and  $\Delta\theta$  can be obtained from the statistical distribution of simulation results. The decreased flexibility of the *trans*-dimer with respect to the monomer is evident from the fact that  $h_T$  is less than  $h_M$ . Movies describing molecular motion of the monomers and dimers are included in Supplementary Information.

the range of allowable values of  $h$ ,  $\Delta\theta$  and  $\Delta\psi$  is significantly reduced, thus resulting in a large entropic penalty for dimerization. By contrast, interdomain mobility is already reduced in *trans*-dimers, such that the additional entropic penalty associated with the *cis*-dimerization of two *trans*-dimers is small in comparison with that between monomers.

We have previously described the process of adherens junction formation as a phase transition between a dilute phase of monomers and *trans*-dimers that diffuse over the surface of a cell, and a condensed lattice composed of *trans*-dimers interacting laterally through a

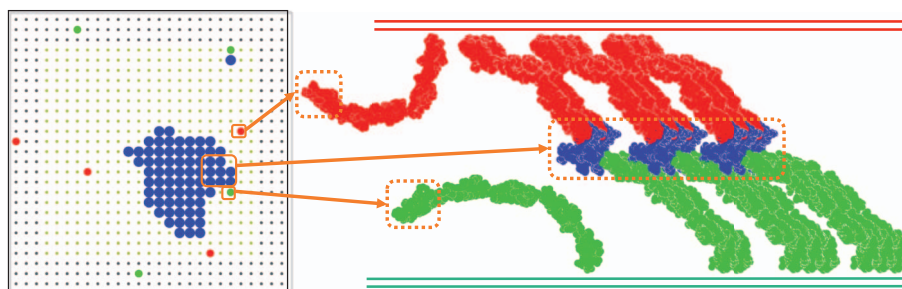
well-defined *cis*-interface<sup>3</sup>. Using lattice simulations, we showed that the formation of a condensed, ordered phase requires *trans*- and *cis*-interactions of sufficient magnitude. The results of such simulations, using the 2D binding affinities reported in Supplementary Table 1, illustrate the formation of well-defined lateral clusters (Fig. 4). Thus, converting the measured 3D cadherin binding affinities into 2D free energies yields interactions of sufficient strength to drive *trans*-dimer formation, and *cis*-interactions between *trans*-dimers of sufficient strength to drive the formation of ordered clusters of these dimers. That is, the values of  $K_d^{(2D)}$  derived here from a combination of experiment, theory and simulation predict that cadherin ectodomains will form junctions, as is observed. By contrast, owing to the one-dimensional nature of *cis*-interactions between monomers (Fig. 1), and because of their small magnitude, monomer oligomerization is negligible<sup>3</sup>.

It is important to note that the treatment we present is based entirely on forces localized to the extracellular region. This is justified for cadherins because junction-like structures form when cytoplasmic regions are deleted<sup>10,17</sup>. However, as has recently been demonstrated for T-cell receptor/major histocompatibility complex interactions, cytoskeletal forces can affect the kinetic and thermodynamic properties of extracellular domains<sup>18</sup>. Thus, although we expect cadherin junction formation *in vivo* to be affected significantly by cytoplasmic involvement, the process is almost certain to depend on the principles of ordered ectodomain assembly uncovered here.

Finally, the concepts and methods introduced in this work should facilitate the analysis of both *trans*- and *cis*-binding interactions between other flexible membrane-bound molecules. For example, chimaeras of CD48 with two or three additional immunoglobulin-like domains are ten times less efficient in adhesion than the wild-type protein, despite having the same binding interface as CD2 (ref. 19). The entropic penalty associated with restricting interdomain motion as a consequence of *trans*-binding provides a simple explanation of these observations and, more generally, offers a mechanism to control binding affinities of membrane-bound receptors that is not available to molecules that are free in solution.

## METHODS SUMMARY

Monte Carlo simulations are carried out in which cadherins domains, each treated as a rigid body described at the level of C $\alpha$  atoms, are allowed to move with respect to the membrane surface through random changes in the three Euler angles,  $\Phi$ ,  $\Theta$  and  $\Psi$ , of the EC5 domain and through motions around the dihedral angles in the hinge regions as indicated in Fig. 3. The angle  $\Phi$  ranges over 360°, whereas  $\Theta$  and  $\Psi$  are restricted to a limited range (0° in one set of simulations and 30° in the



**Figure 4 | Simulation of junction formation.** The lattice in the left panel is a snapshot from a Monte Carlo simulation where cadherin monomers on apposing cells are coloured in red and green, respectively, and *trans*-dimers are coloured blue<sup>3</sup>. A diffusion trap mechanism<sup>3</sup> in which the trap region comprises 20 × 20 lattice sites (yellow/green), in the centre of a 2D lattice of 100 × 100 sites, with periodic boundary conditions, was used in the simulations. *Trans*-dimer formation can take place only in the trap region, as the distance between membranes in the surrounding region is too large to allow *trans*-dimer formation. The cadherins form ordered clusters in the trap region, as indicated. Details of the structure appear in ref. 10. A movie describing the formation of the ordered junction is included in Supplementary Information. The

simulations are carried out using the value of  $K_d^{(2D)}$  (*trans*) for the *trans*-dimerization of E-cadherin (Supplementary Table 1) that is derived from experimental measurements. The total concentration of monomers in each of the two adhering surfaces (either free or *trans*-dimerized) is 1%, whereas the local concentration in the trap region is much higher (18.5%). The corresponding molecular structures of monomers on both cell surfaces, and part of the cluster formed by eight *trans*-dimers, are reconstructed in the right panel from the crystal structure of C-cadherin<sup>15</sup> using the same colour coding. The figure shows the C $\alpha$  backbone with spheres placed on each carbon atom to improve clarity.

other). Motions around the flexible linker regions are described using the elastic network model<sup>20,21</sup>, which defines normal modes along which interdomain motion is allowed. We applied the block normal mode approach<sup>21,22</sup> to partition the structure of the cadherin ectodomain into five rigid blocks, each corresponding to one extracellular domain. The six lowest-frequency modes, each of which describes a collective motion of the entire ectodomain, were used to generate alternative conformations. Fluctuations of the distances between the centres of mass were obtained from molecular dynamics simulations<sup>23</sup> and were used to calibrate the size of the Monte Carlo steps along the normal modes.

In each Monte Carlo step, the EC5 domain was allowed to rotate randomly and the conformation of the whole ectodomain was then changed along one of the normal modes starting with the C-cadherin monomer conformation. For *trans*- and *cis*-dimers, two ectodomains were first placed in conformations generated from the crystal structure of C-cadherin<sup>15</sup>, after which Monte Carlo steps were taken. Two monomers were defined as forming a dimer if the root-mean-square distance obtained from a structural superposition was less than 6 Å, a value determined from molecular dynamics simulations<sup>23</sup> as preserving the dimer interface. Values of  $h_M$ ,  $h_T$ ,  $\Delta\theta_M$ ,  $\Delta\theta_T$ ,  $\Delta\psi_M$  and  $\Delta\psi_T$  were obtained directly from the conformations generated in the Monte Carlo simulations.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 28 December 2010; accepted 6 May 2011.**

- Aplin, A. E., Howe, A. K. & Juliano, R. L. Cell adhesion molecules, signal transduction and cell growth. *Curr. Opin. Cell Biol.* **11**, 737–744 (1999).
- Aricescu, A. R. & Jones, E. Y. Immunoglobulin superfamily cell adhesion molecules: zippers and signals. *Curr. Opin. Cell Biol.* **19**, 543–550 (2007).
- Wu, Y. *et al.* Cooperativity between *trans* and *cis* interactions in cadherin-mediated junction formation. *Proc. Natl Acad. Sci. USA* **107**, 17592–17597 (2010).
- Dustin, M. L., Ferguson, L. M., Chan, P. Y., Springer, T. A. & Golan, D. E. Visualization of CD2 interaction with LFA-3 and determination of the two-dimensional dissociation constant for adhesion receptors in a contact area. *J. Cell Biol.* **132**, 465–474 (1996).
- Dustin, M. L., Bromley, S. K., Davis, M. M. & Zhu, C. Identification of self through two-dimensional chemistry and synapses. *Annu. Rev. Cell Dev. Biol.* **17**, 133–157 (2001).
- Bell, G. I. Models for the specific adhesion of cells to cells. *Science* **200**, 618–627 (1978).
- Bell, G. I., Dembo, M. & Bongrand, P. Cell adhesion. Competition between nonspecific repulsion and specific bonding. *Biophys. J.* **45**, 1051–1064 (1984).
- Chen, C. P., Posy, S., Ben-Shaul, A., Shapiro, L. & Honig, B. H. Specificity of cell-cell adhesion by classical cadherins: critical role for low-affinity dimerization through beta-strand swapping. *Proc. Natl Acad. Sci. USA* **102**, 8531–8536 (2005).
- Patel, S. D. *et al.* Type II cadherin ectodomain structures: implications for classical cadherin specificity. *Cell* **124**, 1255–1268 (2006).
- Harrison, O. J. *et al.* The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure* **19**, 244–256 (2011).
- Katsamba, P. *et al.* Linking molecular affinity and cellular specificity in cadherin-mediated adhesion. *Proc. Natl Acad. Sci. USA* **106**, 11594–11599 (2009).
- Gov, N. S. & Safran, S. A. Red blood cell membrane fluctuations and shape controlled by ATP-induced cytoskeletal defects. *Biophys. J.* **88**, 1859–1874 (2005).
- Zilker, A., Engelhardt, H. & Sackmann, E. Dynamic reflection interference contrast (RIC-) microscopy: a new method to study surface excitations of cells and to measure membrane bending elastic moduli. *J. Phys.* **48**, 2139–2151 (1987).
- Hill, T. L. *An Introduction to Statistical Thermodynamics* 147–176 (Dover, 1987).
- Boggon, T. J. *et al.* C-cadherin ectodomain structure and implications for cell adhesion mechanisms. *Science* **296**, 1308–1313 (2002).
- Dustin, M. L. *et al.* Low affinity interaction of human or rat T cell adhesion molecule CD2 with its ligand aligns adhering membranes to achieve high physiological affinity. *J. Biol. Chem.* **272**, 30889–30898 (1997).
- Hong, S., Troyanovsky, R. B. & Troyanovsky, S. M. Spontaneous assembly and active disassembly balance adherens junction homeostasis. *Proc. Natl Acad. Sci. USA* **107**, 3528–3533 (2010).
- Huppa, J. B. *et al.* TCR-peptide-MHC interactions in situ show accelerated kinetics and increased affinity. *Nature* **463**, 963–967 (2010).
- Milstein, O. *et al.* Nanoscale increases in CD2–CD48-mediated intermembrane spacing decrease adhesion and reorganize the immunological synapse. *J. Biol. Chem.* **283**, 34414–34422 (2008).
- Atilgan, A. R. *et al.* Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **80**, 505–515 (2001).
- Li, G. H. & Cui, Q. A coarse-grained normal mode approach for macromolecules: an efficient implementation and application to  $\text{Ca}^{2+}$ -ATPase. *Biophys. J.* **83**, 2457–2474 (2002).
- Tama, F., Gadea, F. X., Marques, O. & Sanejouand, Y. H. Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins* **41**, 1–7 (2000).
- Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported by National Science Foundation grant MCB-0918535 (to B.H.) and National Institutes of Health grant R01 GM062270-07 (to L.S.). The financial support of the US-Israel Binational Science Foundation (grant no. 2006-401, to A.B.-S., B.H. and L.S.) and the Israel Science Foundation (ISF 1448/10 and 695/06) (to A.B.-S.) is acknowledged. We thank E. Sackmann for an email exchange concerning membrane fluctuations.

**Author Contributions** Y.W., J.V., L.S., B.H. and A.B.-S. designed the research; Y.W. performed the multiscale simulations; J.V. carried out the all-atom molecular dynamics simulations; Y.W., B.H. and A.B.-S. analysed the data; Y.W., A.B.-S. and B.H. contributed analytic tools; and Y.W., L.S., B.H. and A.B.-S. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to A.B.-S. ([abs@fh.huji.ac.il](mailto:abs@fh.huji.ac.il)) or B.H. ([bh6@columbia.edu](mailto:bh6@columbia.edu)).

## METHODS

**Modelling intramolecular conformational changes.** We used the elastic network model<sup>20</sup> to define normal modes to be used in the Monte Carlo simulations. This model is based on the approximation that molecular vibrations near an equilibrium conformation can be determined by a coarse-grained harmonic potential

$$V = \frac{\gamma}{2} \sum_{ij} \sigma_{ij} \left( |\mathbf{r}_{ij}| - |\mathbf{r}_{ij}^0| \right)^2$$

$$\sigma_{ij} = \begin{cases} 1, & |\mathbf{r}_{ij}^0| \leq r_c \\ 0, & |\mathbf{r}_{ij}^0| > r_c \end{cases}$$

where  $|\mathbf{r}_{ij}|$  and  $|\mathbf{r}_{ij}^0|$  are respectively the instantaneous and equilibrium values of the distance between C $\alpha$  atoms  $i$  and  $j$ ,  $\gamma$  is the uniform force constant and the cut-off value,  $r_c$ , is set as 13 Å. We applied the block normal mode approach<sup>21,22</sup> to partition the structure of each cadherin ectodomain into five blocks, each corresponding to one extracellular domain. A C $\alpha$  representation of the native structure of C-cadherin<sup>15</sup> (Protein Data Bank ID, 1L3W) was used as an initial model. We chose the six lowest-frequency modes, all of which describe collective motions of the entire ectodomain. Amplitudes of motion along the direction defined by each normal mode were obtained from the following procedure.

**Calibrating the amplitude of interdomain motions.** A 40-ns all-atom molecular dynamics simulation of the EC1 and EC2 domains was carried out in explicit solvent using GROMACS<sup>23</sup>. For structures generated along the simulation trajectory, the coordinates of the EC2 domain were fixed and then the distance between the centre of mass of the EC1 domain in each simulation step and its centre of mass in the initial conformation was calculated. Supplementary Fig. 1a plots this distance versus simulation time. After transforming this fluctuation profile into a frequency-like histogram, a Gaussian-like distribution with a range of about  $\pm 8$  Å is obtained (Supplementary Fig. 1b).

To relate this distance fluctuation to corresponding motions along the six normal modes, we performed a series of Monte Carlo tests where we started with the crystal structure and generated a series of different conformations by taking a random step, smaller than some pre-chosen cut-off, along any one of the six eigenvectors. The centre of mass of EC5 was kept fixed, so that no 2D diffusion occurred in this stage. All conformations generated with a single Monte Carlo test have the same cut-off value. A total of 20 cut-offs were tried with the goal of finding a value (and a corresponding step size) that would reproduce, as closely as possible, the distribution of EC1–EC2 distances obtained from the molecular dynamics simulations. To this end, all normal-mode-generated structural models obtained using the same cut-off were aligned to one another by superimposing their EC2 domains, yielding an ensemble of EC1 domain positions. By calculating the distance between the EC1 domain centre of mass in each structural model and the EC1 domain centre of mass in the crystal structure, a histogram of this distance distribution was generated for each cut-off. The Monte Carlo step size along normal modes was defined so that the range of normal-mode-generated distance

distributions was as close as possible to 8 Å (Supplementary Fig. 1c), the value generated by the all-atom simulations.

**Estimating geometric variables of EC1 domain fluctuations.** Different conformations of a cadherin monomer were generated with a Monte Carlo simulation using the normal modes and step sizes derived from the methods described above. In each step of the simulation, the EC5 domain is first allowed to randomly rotate within a small interval in Euclidean space  $\Phi$ – $\Theta$ – $\Psi$ , as shown in the upper left panel of Fig. 3. Then the conformation of the entire ectodomain is changed, in a positive or negative direction, along one of the six normal modes using a step size chosen randomly from the range of values that produce the distribution shown in Supplementary Fig. 1c. A large number of structures are generated in this way as shown in Fig. 3. The fluctuations of the centre of mass of the EC1 domains along the  $z$  axis, as well as fluctuations in the Euler angles, are obtained directly from a straightforward geometric analysis of these structures:  $h_M$  is defined as twice the standard deviation of the distribution of the centre of mass of the EC1 domain along the  $z$  axis, and the distributions of the Euler angles, defined by  $\Delta\theta_M$  and  $\Delta\psi_M$ , are obtained in the same way. Results are shown in Supplementary Table 1 and Supplementary Fig. 2.

For a *trans*-dimer, two cadherin ectodomains are initially placed facing each other, as in the native structure of the *trans*-dimer of C-cadherin<sup>15</sup>. Then the conformation of each monomer is randomly modified using the algorithm described above for monomers. Intermolecular clashes are checked after each Monte Carlo step. If there is no severe intermolecular clash, and the distance between the centres of mass of the two EC1 domains is less than 50 Å, the root mean-square distance (r.m.s.d.) of the EC1 domain pair relative to the native strand-swapped-dimer is calculated. Two cadherin monomers are defined as forming a *trans*-dimer if the EC1 r.m.s.d. is less than 6 Å. This cut-off value was determined from all-atom molecular dynamics simulations of a *trans*-dimer formed by C-cadherin EC1–EC2 domains. During the molecular dynamics simulations, the dimer structure deviated from the initial conformation as, for example, can be seen in Supplementary Fig. 3, which shows a structural superposition of two conformations, one in the initial state (red) and one chosen from the middle of the simulation (green). The C $\alpha$  r.m.s.d. between the two EC1 domains is about 4 Å. Supplementary Fig. 3b shows two independent trajectories of EC1 r.m.s.d. fluctuations obtained from the GROMACS<sup>23</sup> molecular dynamics simulations. As can be seen in the figure, the r.m.s.d. fluctuations from the native structure are within the range of 6 Å in both simulations.

The range of EC1 domain fluctuations along the  $z$  axis,  $h_T$ , and the rotational distribution,  $\Delta\theta_T$  and  $\Delta\psi_T$ , were determined and are reported in Supplementary Table 1 and Supplementary Fig. 4. The procedure was the same as used for monomers except that the fluctuations of each EC1 domain in a *trans*-dimer were included separately in the distribution. The same procedure was used for *cis*-dimers as well, but in this case two cadherin extracellular domains were placed on the same surface and their initial orientations were based on the *cis*-interface taken from the C-cadherin crystal structure<sup>15</sup>. The range of domain fluctuations along the  $z$  axis,  $h_C$ , and the rotational distribution,  $\Delta\theta_C$  and  $\Delta\psi_C$ , were determined and are reported in Supplementary Table 1 and Supplementary Fig. 5.



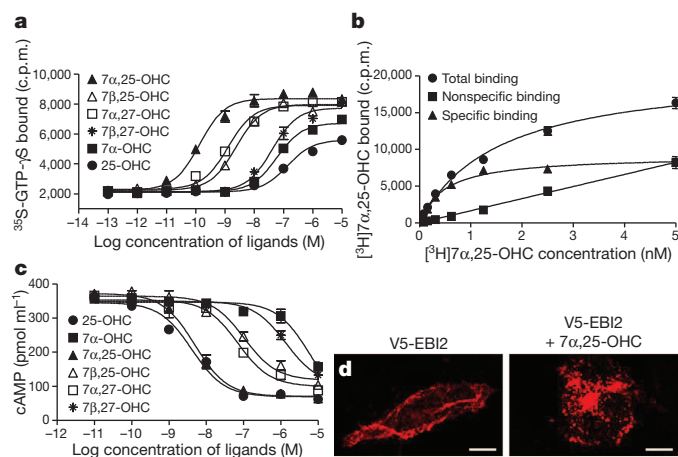
# Oxysterols direct B-cell migration through EBI2

Changlu Liu<sup>1</sup>, Xia V. Yang<sup>1</sup>, Jiejun Wu<sup>1</sup>, Chester Kuei<sup>1</sup>, Neelakandha S. Mani<sup>1</sup>, Li Zhang<sup>1</sup>, Jingxue Yu<sup>1</sup>, Steven W. Sutton<sup>1</sup>, Ning Qin<sup>1†</sup>, Homayon Banie<sup>1</sup>, Lars Karlsson<sup>1†</sup>, Siquan Sun<sup>1</sup> & Timothy W. Lovenberg<sup>1</sup>

EBI2 (also called GPR183) is an orphan G-protein-coupled receptor that is highly expressed in spleen and upregulated upon Epstein–Barr-virus infection<sup>1</sup>. Recent studies indicated that this receptor controls follicular B-cell migration and T-cell-dependent antibody production<sup>2–6</sup>. Oxysterols elicit profound effects on immune and inflammatory responses as well as on cholesterol metabolism<sup>7–9</sup>. The biological effects of oxysterols have largely been credited to the activation of nuclear hormone receptors<sup>10,11</sup>. Here we isolate oxysterols from porcine spleen extracts and show that they are endogenous ligands for EBI2. The most potent ligand and activator is 7 $\alpha$ ,25-dihydroxycholesterol (OHC), with a dissociation constant of 450 pM for EBI2. *In vitro*, 7 $\alpha$ ,25-OHC stimulated the migration of EBI2-expressing mouse B and T cells with half-maximum effective concentration values around 500 pM, but had no effect on EBI2-deficient cells. *In vivo*, EBI2-deficient B cells or normal B cells desensitized by 7 $\alpha$ ,25-OHC pre-treatment showed reduced homing to follicular areas of the spleen. Blocking the synthesis of 7 $\alpha$ ,25-OHC *in vivo* with clotrimazole, a CYP7B1 inhibitor, reduced the content of 7 $\alpha$ ,25-OHC in the mouse spleen and promoted the migration of adoptively transferred pre-activated B cells to the T/B boundary (the boundary between the T-zone and B-zone in the spleen follicle), mimicking the phenotype of pre-activated B cells from EBI2-deficient mice. Our results show an unexpected causal link between EBI2, an orphan G-protein-coupled receptor controlling B-cell migration, and the known immunological effects of certain oxysterols, thus uncovering a previously unknown role for this class of molecules.

Oxysterols are oxygenated derivatives of cholesterol synthesized through enzymatic<sup>12</sup> or non-enzymatic<sup>13</sup> processes. Oxysterols exert profound biological effects in cholesterol and fatty-acid metabolism, immune regulation, and possibly neurodegenerative diseases<sup>7,8</sup>. EBI2 is an orphan G-protein-coupled receptor (GPCR) that was first identified as a gene highly upregulated by cellular infection with Epstein–Barr virus<sup>1</sup>. Using tissue extracts, we identified putative EBI2 ligand activity in rat spleen (Supplementary Fig. 1a) and purified EBI2 ligands from porcine spleen through organic extractions followed by chromatography (Supplementary Fig. 1b). Gas chromatography-mass spectrometry (GCMS) showed that the porcine spleen active peak B (Supplementary Fig. 1b) contains 7 $\alpha$ -hydroxycholesterol (7 $\alpha$ -OHC) and 7 $\beta$ -OHC (Supplementary Fig. 1c), which was further confirmed by NMR (Supplementary Fig. 1d, e). We then tested more than 30 oxysterols for the activation of EBI2. 7 $\alpha$ -OHC, 7 $\beta$ -OHC, 25-OHC, 27-OHC, 7 $\alpha$ ,27-OHC and 7 $\beta$ ,27-OHC were found to stimulate <sup>35</sup>S-GTP- $\gamma$ S incorporation in EBI2-expressing membranes with various potencies and efficacies (Fig. 1a and Supplementary Table 1), but not in control cell membranes. Structurally related compounds including vitamin D derivatives, steroids and bile acids were found to be inactive (Supplementary Table 2). Among the active ligands, 7 $\alpha$ ,27-OHC showed the highest potency (half-maximum effective concentration (EC<sub>50</sub>) = 1.3 nM). 7 $\beta$ ,27-OHC showed an EC<sub>50</sub> value of 51 nM, whereas 7 $\alpha$ -OHC, 7 $\beta$ -OHC, 25-OHC and 27-OHC were all less potent and appeared to be partial agonists. We noticed that 7 $\alpha$ -OHC was

more potent and efficacious than 7 $\beta$ -OHC, whereas 25-OHC was more potent and efficacious than 27-OHC. Therefore, we hypothesized that 7 $\alpha$ ,25-OHC, which is a known endogenous oxysterol converted from 25-OHC by CYP7B1 (ref. 14), might be even more potent than 7 $\alpha$ ,27-OHC. We synthesized 7 $\alpha$ ,25-OHC and confirmed our prediction that it is highly potent (Fig. 1a) at activating EBI2 (EC<sub>50</sub> = 140 pM), whereas 7 $\beta$ ,25-OHC was less potent (EC<sub>50</sub> = of 2.1 nM). Comparing the retention times of the active oxysterols using the same chromatography method that initially identified the ligand activities from the porcine spleen extracts, we found that 25-OHC, 7 $\alpha$ -OHC/7 $\beta$ -OHC and 7 $\alpha$ ,25-OHC co-migrate with the activity peaks A, B and C (Supplementary Fig. 1b), respectively. These findings indicated that the active components in peak A and peak C are likely to be 25-OHC and 7 $\alpha$ ,25-OHC. Recent studies indicate that there are distinct cholesterol binding sites on GPCRs<sup>15</sup>. Here we show that oxysterols are specific ligands for EBI2 because they lack ligand activity for other GPCRs (Supplementary Fig. 2). <sup>3</sup>H-labelled 7 $\alpha$ ,25-OHC was generated by converting <sup>3</sup>H-labelled 25-OHC in cultured COS7 cells expressing recombinant CYP7B1 enzyme and then purified by high-performance liquid chromatography (HPLC) (Supplementary Fig. 3a). Radioligand binding studies using this tracer showed specific binding of [<sup>3</sup>H]7 $\alpha$ ,25-OHC in EBI2-expressing cells but not in mock-transfected cells (Supplementary



**Figure 1 | Identification and pharmacological characterization of oxysterols as ligands for EBI2.** **a**, Characterization of oxysterols as agonists for EBI2. Various concentrations of different oxysterols were used to stimulate <sup>35</sup>S-GTP- $\gamma$ S incorporation in EBI2-expressing cell membranes. c.p.m., counts per min. **b**, Saturation binding study of EBI2 using [<sup>3</sup>H]-labelled 7 $\alpha$ ,25-OHC as the radioligand. **c**, EBI2 ligands suppress forskolin-induced cAMP accumulation in EBI2-expressing cells. **d**, 7 $\alpha$ ,25-OHC stimulates EBI2 internalization. Human EBI2 with a V5 tag at the N terminus was expressed in CHO cells. Anti-V5 antibody (primary antibody) and Cy3-labelled goat-anti-mouse IgG (secondary antibody) were used to monitor the receptor localization, which was visualized by confocal microscopy. No staining was observed in CHO cells without recombinant EBI2 expression. Scale bars, 10  $\mu$ m. The results in **a–c** represent mean  $\pm$  s.e.m.,  $n = 3$ .

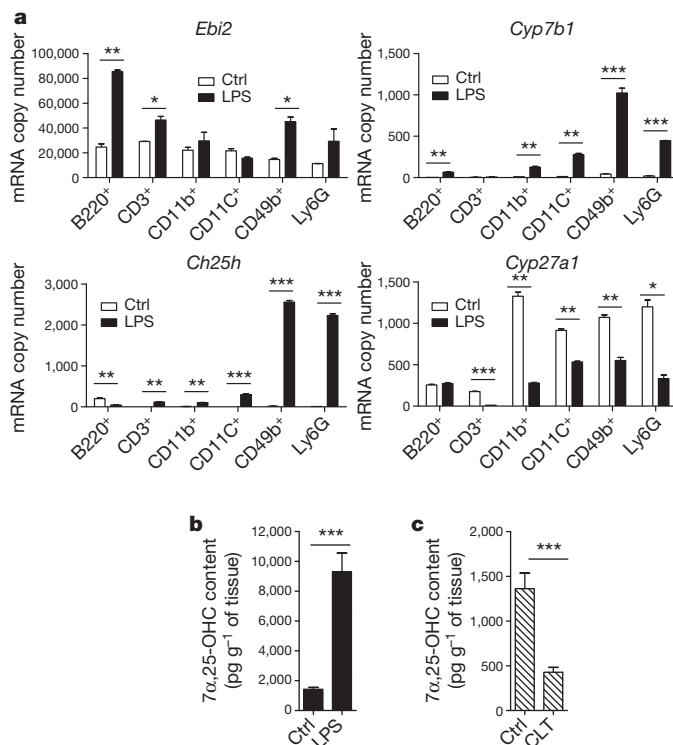
<sup>1</sup>Johnson & Johnson Pharmaceutical Research & Development, L.L.C., 3210 Merryfield Row, San Diego, California 92121, USA. <sup>†</sup>Present addresses: Roche R&D Center (China) Ltd., 720 Cai Lun Road, Building 5, Zhangjiang High-Tech Park, Pudong, Shanghai 201203, China (N.Q.); Regulus Therapeutics, 3545 John Hopkins Court, San Diego, California 92121, USA (L.K.).

Fig. 3b). Saturation binding analysis revealed that [ $^3\text{H}$ ]7 $\alpha$ ,25-OHC binds to EBI2 with a dissociation constant value of 450 pM (Fig. 1b). In a competition binding study, 7 $\alpha$ ,25-OHC was able to compete for binding with a very high potency (half-maximum inhibitory concentration ( $\text{IC}_{50}$ ) = 242 pM), whereas 7 $\alpha$ ,27-OHC, 7 $\beta$ ,25-OHC, 7 $\beta$ ,27-OHC, 25-OHC and 7 $\alpha$ -OHC showed lower potencies with  $\text{IC}_{50}$  values of 799 pM, 4.3 nM, 337 nM, 1.26  $\mu\text{M}$  and 1.32  $\mu\text{M}$ , respectively (Supplementary Fig. 3c). The observed binding affinity correlated with the functional agonist potency for all ligands tested.

7 $\alpha$ ,25-OHC stimulates GTP- $\gamma\text{S}$  binding in cells co-expressing EBI2 and  $\text{G}_{\text{o}2}$  protein, indicating that EBI2 is coupled to  $\text{G}_{\text{i/o}}$  G proteins. Pertussis toxin treatment suppressed EBI2 activity, thus supporting this hypothesis (Supplementary Fig. 4a). To investigate further whether EBI2 suppresses cAMP accumulation, a SK-N-MC/CRE- $\beta$ -galactosidase cell line<sup>16</sup> stably expressing human EBI2 was evaluated. 7 $\alpha$ ,25-OHC dose-dependently suppressed forskolin-induced cAMP accumulation in EBI2-expressing cells (Fig. 1c), but not in control cells (Supplementary Fig. 4b), with an  $\text{IC}_{50}$  value of 2 nM. Compared with 7 $\alpha$ ,25-OHC, other active oxysterols, except 25-OHC, showed lower potencies with a rank order similar to that observed in the GTP- $\gamma\text{S}$  binding assay. 25-OHC, which displayed a low potency in the GTP- $\gamma\text{S}$  binding assay and radioligand binding assay, showed surprisingly similar high potency to that of 7 $\alpha$ ,25-OHC in direct cAMP measurement and a  $\beta$ -gal-linked cAMP reporter assay (Fig. 1c and Supplementary Fig. 4c, d). Because compounds were cultured with live cells for an extended period of time in both cAMP accumulation and  $\beta$ -gal assays, we reasoned that 25-OHC might have been converted into 7 $\alpha$ ,25-OHC by CYP7B1 endogenously expressed by the host cells. Supporting this hypothesis, clotrimazole, a CYP7B1 inhibitor<sup>14</sup>, inhibits 25-OHC activity but not 7 $\alpha$ ,25-OHC activity (Supplementary Fig. 5). GCMS analysis showed that, upon culturing with SK-N-MC cells, 25-OHC was converted into 7 $\alpha$ ,25-OHC (Supplementary Fig. 6).

To study the membrane trafficking of EBI2, we engineered an aminoterminal V5-tagged receptor, which demonstrated almost identical pharmacology in the GTP- $\gamma\text{S}$  binding assay compared with wild-type EBI2 (Supplementary Fig. 7). Surface V5–EBI2 protein was first labelled with anti-V5 antibody in cultured cells, followed by stimulation with 7 $\alpha$ ,25-OHC. Without stimulation, V5–EBI2 had a very typical cell surface distribution. After stimulation, staining of V5–EBI2 became very prominent in the intracellular organelles, indicating ligand-stimulated receptor internalization of EBI2 (Fig. 1d). Quantitative enzyme-linked immunosorbent assay analysis showed that after 7 $\alpha$ ,25-OHC stimulation, the cell surface staining of EBI2 had been markedly reduced (Supplementary Fig. 8). The receptor internalization studies indicate that EBI2 may interact with  $\beta$ -arrestin, which is confirmed by a  $\beta$ -arrestin recruitment study (Supplementary Fig. 9).

Among the active oxysterols, 7 $\alpha$ ,25-OHC and 7 $\alpha$ ,27-OHC are potential endogenous ligands of EBI2 based on their higher *in vitro* potency. We first examined the expression of EBI2 and the enzymes for 7 $\alpha$ ,25-OHC and 7 $\alpha$ ,27-OHC biosynthesis including CH25H (ref. 17), CYP27A1 (ref. 18) and CYP7B1 (ref. 14) in purified subtypes of mouse splenocytes. CYP7B1 is particularly critical because it is required for the final stage synthesis of 7 $\alpha$ ,25-OHC and 7 $\alpha$ ,27-OHC. CH25H and CYP7B1 have previously been shown to be highly upregulated in macrophages<sup>19</sup> after infection or lipopolysaccharide (LPS) stimulation<sup>20,21</sup>. We found that in splenocytes from naive mice, the expression of CYP7B1 and CYP27A1 is higher in macrophages, dendritic cells, natural killer (NK) cells and neutrophils, whereas B cells express higher levels of CH25H. However, upon LPS activation (Fig. 2a), in addition to the upregulation of EBI2 expression, the expression of CYP7B1 and CH25H is markedly upregulated in macrophages, dendritic cells, and especially in NK cells and neutrophils, whereas CYP27A1 expression is downregulated in all cell types. The synchronized upregulation of EBI2, CH25H and CYP7B1 expression by LPS indicates that 7 $\alpha$ ,25-OHC is probably the endogenous ligand for EBI2. In contrast, the downregulation of CYP27A1 expression by LPS



**Figure 2 | Regulation of EBI ligand synthesis in mouse spleen.**

**a**, Quantitative RT-PCR analysis of mRNA expression for *Ebi2*, *Cyp7b1*, *Ch25h* and *Cyp27a1* in various types of purified splenocytes, B cells (B220<sup>+</sup>), T cells (CD4<sup>+</sup> or CD8<sup>+</sup>), macrophages (CD11b<sup>+</sup>), dendritic cells (CD11c<sup>+</sup>), natural killer cells (CD49b<sup>+</sup>) and neutrophils (Ly6G<sup>+</sup>) from mice either treated with saline (Ctrl) or LPS (LPS). The mRNA expression levels of *Ebi2*, *Cyp7b1*, *Ch25h* and *Cyp27a1* were normalized to the expression levels of  $\beta$ -actin mRNA and expressed as copy number per 10<sup>6</sup> copies of  $\beta$ -actin mRNA. The results represent mean  $\pm$  s.e.m., \*\*\* $P$  < 0.0001, \*\* $P$  < 0.01, \* $P$  < 0.05, unpaired, two-tailed  $t$ -test,  $n$  = 2. **b**, Quantitative analysis of 7 $\alpha$ ,25-OHC in mouse spleens treated with saline (Ctrl) or LPS (LPS). \*\*\* $P$  < 0.0001, unpaired, two-tailed  $t$ -test,  $n$  = 8. **c**, Spleen content of 7 $\alpha$ ,25-OHC in mice treated with vehicle (Ctrl) or clotrimazole (CLT). \*\*\* $P$  < 0.0001, unpaired, two-tailed  $t$ -test,  $n$  = 8. Error bars for **b** and **c** represent mean  $\pm$  s.e.m.

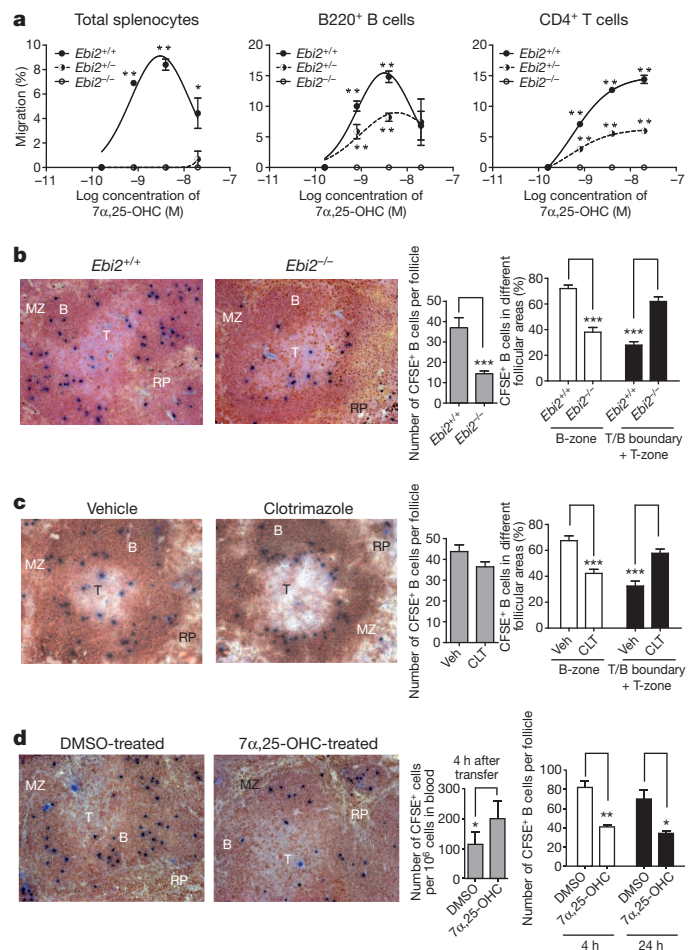
indicates that 7 $\alpha$ ,27-OHC may not be the primary EBI2 ligand. To examine the production of EBI2-activating oxysterols *in vivo*, we extracted total oxysterols from freshly dissected mouse spleens followed by HPLC fractionation and quantified ligand activities using EBI2-expressing membrane in the GTP- $\gamma\text{S}$  binding assay. The only apparent ligand activity observed was corresponding to the 7 $\alpha$ ,25-OHC-containing fraction (Supplementary Fig. 10), which was further confirmed by GCMS analysis. Because the fractions were reconstituted in assay volumes comparable to the original tissue volumes, the results indicated that the content of 7 $\alpha$ ,25-OHC in mouse spleen is sufficient for activation of EBI2 *in vivo*. Consistent with an increased expression in CH25H and CYP7B1, we observed that the content of 7 $\alpha$ ,25-OHC was markedly upregulated by LPS challenge (Fig. 2b). In control mice, 7 $\alpha$ ,25-OHC has an average content of  $1.39 \pm 0.18 \text{ ng g}^{-1}$  of spleen tissue and its content increased to  $9.38 \pm 1.31 \text{ ng g}^{-1}$  of spleen tissue after LPS treatment. These results strongly suggest that under the normal physiological conditions and particularly after LPS challenge, 7 $\alpha$ ,25-OHC could reach the concentrations capable of activating EBI2. Compared with 7 $\alpha$ ,25-OHC, the activity corresponding to 7 $\alpha$ ,27-OHC was not detected neither at the normal condition nor after LPS stimulation. Considering that 7 $\alpha$ ,27-OHC also has lower potency and efficacy, it is unlikely that 7 $\alpha$ ,27-OHC serves as the primary activator for EBI2 in the spleen. However, we cannot exclude the possibility that under some special conditions or in other tissues where 7 $\alpha$ ,27-OHC may reach sufficient levels to activate EBI2. Clotrimazole inhibits



7 $\alpha$ ,25-OHC synthesis, which has been reported previously<sup>14</sup> and is supported by our *in vitro* results (Supplementary Fig. 5). When administered to mice *in vivo*, clotrimazole decreased the content of 7 $\alpha$ ,25-OHC in the spleen (Fig. 2c). In vehicle-treated mice, the content of 7 $\alpha$ ,25-OHC is about  $1.41 \pm 0.16 \text{ ng g}^{-1}$  of spleen tissue. After treatment with clotrimazole, the content of 7 $\alpha$ ,25-OHC was reduced to  $0.43 \pm 0.06 \text{ ng g}^{-1}$  of tissue. These results indicate that clotrimazole can be used as a pharmacological tool to downregulate the content of 7 $\alpha$ ,25-OHC and therefore may inhibit the activation of EBI2 *in vivo*.

It was recently reported that EBI2 seemed to have an essential role in guiding activated B cells to inter/outer follicles and to extra-follicular areas<sup>2-6</sup>. Our finding of oxysterols as ligands for EBI2 indicates that certain oxysterols might be chemotactic factors for B cells. Recombinant mouse EBI2 responded to the active oxysterols with very similar pharmacology to that of the human receptor (Supplementary Fig. 11). In an *in vitro* migration assay, we observed that 7 $\alpha$ ,25-OHC dose-dependently stimulated migration of *in vitro* LPS-activated spleen B cells and anti-CD3/CD28-activated CD4<sup>+</sup> T cells (Fig. 3a) as well as non-activated (although to a lesser extent) and *in vivo* LPS-activated B cells, CD4<sup>+</sup> and CD8<sup>+</sup> T cells, and dendritic cells (Supplementary Fig. 12a, b). This was supported by the finding that T cells and dendritic cells, like B cells, also express EBI2 (Fig. 2a). Although macrophages, NK cells and neutrophils also express EBI2, little 7 $\alpha$ ,25-OHC-induced migration *in vitro* was observed for these cells. 7 $\alpha$ ,25-OHC exhibited a bell-shaped dose-response curve in inducing B-cell migration (Fig. 3a), indicating ligand-induced receptor desensitization, possibly via receptor internalization as described above. 7 $\alpha$ ,27-OHC is also active in inducing splenocyte migration (albeit with reduced potency compared to 7 $\alpha$ ,25-OHC). 7 $\beta$ ,25-OHC, 7 $\beta$ ,27-OHC, 7 $\alpha$ -OHC, 7 $\beta$ -OHC and 25-OHC were essentially inactive at up to 100 nM (Supplementary Fig. 12c), indicating that these oxysterols are unlikely to be physiological EBI2 ligands. These results are consistent with our observation in a  $\beta$ -arrestin recruitment assay using recombinant EBI2-expressing cells (Supplementary Fig. 12d). Perhaps the EBI2-directed cell migration is related to the  $\beta$ -arrestin activation pathway. To confirm that 7 $\alpha$ ,25-OHC-induced cell migration is EBI2-dependent, we evaluated its effects in EBI2-deficient mice, in which the *Ebi2* gene was disrupted by replacing the coding region 45–646 (602 bp) with a Neo cassette (7,780 bp) (Supplementary Fig. 13). The results showed that unseparated splenocytes, B cells and T cells from EBI2-deficient mice (*Ebi2*<sup>-/-</sup>) were completely unresponsive to 7 $\alpha$ ,25-OHC stimulation. Notably, B and T cells from heterozygous mice (*Ebi2*<sup>+/-</sup>) showed a reduced response, at about 50% of that found with wild-type cells, in the migration assay (Fig. 3a).

CXCL13/CXCR5, CXCL12/CXCR4 and CCL21 (or CCL19)/CCR7 are three sets of chemokines/receptors known to be critical in guiding migration of naive and active B cells in secondary lymph organs. Whereas CXCL13/CXCR5 drive B-cell follicle and germinal centre localization, CXCL12/CXCR4 direct early plasmablasts to the extra-follicular area and CCL21 (or CCL19)/CCR7 is believed to recruit activated B cells to the T/B boundary<sup>3</sup>. It has been suggested that EBI2 controls differentiation of plasmablasts by promoting migration of activated B cells to inter/outer follicular areas but not germinal centres<sup>2,4</sup>. Previous adoptive transfer studies<sup>2</sup> have shown that EBI2-deficient B cells (24 h after the transfer) mainly localize at the centre of follicles whereas wild-type B cells are found throughout the follicular area. In our adoptive transfer experiments with pre-activated B cells, we found that there was a significant decrease in homing of activated EBI2-deficient B cells to spleen follicles at 4 h (Fig. 3b), but not at 24 h (consistent with previous findings<sup>2</sup>) (Supplementary Fig. 14a), after transfer, indicating that EBI2 may also have a role in B-cell homing to follicles. In addition, owing to the transient upregulation of CCR7 in B cells after activation<sup>22</sup>, following adoptive transfer, some of the pre-activated B cells from wild-type mice were found to migrate to the T/B boundary. However, significant numbers of transferred wild-type B cells were localized throughout other areas of the follicle. In contrast,



**Figure 3 | 7 $\alpha$ ,25-OHC stimulates B-cell and T-cell migration.** **a**, *In vitro* migration studies of activated splenocytes from wild-type (*Ebi2*<sup>+/+</sup>), EBI2-deficient (*Ebi2*<sup>-/-</sup>), or heterozygous (*Ebi2*<sup>+/-</sup>) mice. Starting cells and migrated cells were analysed by FACS. Cell migrations are expressed as percentages of starting total splenocytes, B220<sup>+</sup>, or CD4<sup>+</sup> cells, respectively. Results represent mean  $\pm$  s.e.m.,  $n = 3$ . **b**, EBI2-deficient B cells showed reduced homing to follicles and localized to T/B boundary 4 h after adoptive transfer. *In vitro* activated B cells were labelled with carboxyfluorescein succinimidyl ester (CFSE) and adoptively transferred (intravenously) into normal recipient mice ( $n = 3$ ). Spleen cryo-sections were stained with anti-CFSE (transferred B cells) and anti-IgD/anti-B220 (follicles). The total numbers of CFSE<sup>+</sup> B cells found in representative follicles ( $n = 9-12$ ) as well as in B-zone or T/B boundary (including T-zone) were counted (as shown in Supplementary Fig. 17) and the percentage calculated. **c**, Clotrimazole treatment promotes the adoptively transferred pre-activated B cells to localize at the T/B boundary 4 h after adoptive transfer. The total numbers of CFSE<sup>+</sup> B cells found in representative follicles ( $n = 6-9$ ) as well as in B-zone or T/B boundary (including T-zone) were counted and the percentage calculated.  $n = 3$  for recipient mice, pre-treated with vehicle or clotrimazole (100 mg kg<sup>-1</sup>, dosed twice orally). **d**, Pre-treatment of activated B cells with 7 $\alpha$ ,25-OHC reduces homing of B cells to follicle. *In vitro* activated B cells were treated either with vehicle (DMSO) or 1  $\mu$ M of 7 $\alpha$ ,25-OHC before adoptive transfer into normal recipient mice ( $n = 3$ ). At 4 h or 24 h after transfer, both blood and spleens were harvested and analysed for CFSE<sup>+</sup> B cells by FACS or immunohistochemistry. Six–nine representative follicles were counted. B, B zone; MZ, marginal zone; RP, red pulp; T, T zone. \*\*\* $P < 0.0001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , unpaired, two-tailed *t*-test.

adoptively transferred pre-activated B cells from EBI2-deficient mice resulted in a much greater percentage of B cells localized at the T/B boundary (some B cells were even found in the T-zone) at 4 h (Fig. 3b), but not after 24 h (Supplementary Fig. 14a), which strongly suggests that EBI2 has a role in antagonizing CCR7-driven migration of newly activated B cells to the T/B boundary. To address the question of the



endogenous role of  $7\alpha,25$ -OHC, we blocked its synthesis *in vivo* by dosing recipient mice with clotrimazole (Fig. 2c) to see if we could produce an *in vivo* B-cell migration result similar to that using EBI2-deficient B cells. Notably, pre-activated wild-type B cells adoptively transferred into the clotrimazole-treated mice were found mainly localized at the T/B boundary at 4 h after transfer (Fig. 3c), but not after 24 h (Supplementary Fig. 14b), whereas in the vehicle-treated control mice, adoptively transferred B cells are more evenly distributed within follicles. The results strongly suggest that  $7\alpha,25$ -OHC is an endogenous ligand for EBI2 and the physiological function of this ligand/receptor includes a role in driving B-cell migration *in vivo*.

Although EBI2-deficient B cells were reported to express similar levels of CXCR4, CXCR5 and CCR7, and showed no difference in migration induced by the corresponding chemokines<sup>4</sup>, we reasoned that these chemokines may functionally antagonize  $7\alpha,25$ -OHC-induced migration, and pre-activation of EBI2 with  $7\alpha,25$ -OHC may thus affect chemokine-induced migration. Supporting this hypothesis, CXCL13, CXCL12 and CCL21 were found to inhibit  $7\alpha,25$ -OHC-induced migration of B cells *in vitro* (Supplementary Fig. 15a). Furthermore, pre-treatment of activated B cells with  $7\alpha,25$ -OHC also reduced migration induced by CXCL13, CXCL12, CCL21, or  $7\alpha,25$ -OHC *in vitro* (Supplementary Fig. 15b). The desensitization of chemokine receptors by the EBI2 ligand is probably a result of heterologous desensitization of GPCRs<sup>23</sup>. However, this apparent  $7\alpha,25$ -OHC-induced desensitization to chemokines was not due to downregulation of the cell surface chemokine receptors as assayed by fluorescence-activated cell sorting (FACS) (Supplementary Fig. 15c). The functional desensitization of EBI2 and chemokine receptors indicates that pre-activation of EBI2 *in vitro* may inhibit B cells homing to follicles *in vivo*. Consistent with this hypothesis, B cells pre-treated *in vitro* with  $7\alpha,25$ -OHC, after adoptive transfer into normal recipient mice, were found to show reduced homing to follicular areas in spleen (Fig. 3d). The reduction of B-cell homing to follicles after  $7\alpha,25$ -OHC treatment is not due to decreased cell viability or that fewer cells were transferred, because more  $7\alpha,25$ -OHC pre-treated B cells were found in blood samples (Fig. 3d) compared with B cells treated with vehicle. Additional *in vitro* experiments also showed that  $7\alpha,25$ -OHC does not affect the cell viability (Supplementary Fig. 16a). These results indicate that exposure to  $7\alpha,25$ -OHC could inhibit B-cell migration induced by physiologically relevant levels of chemokines *in vivo* in a normal host spleen.

Recent studies in EBI2-deficient mice indicated that this receptor is also involved in T-cell-dependent antibody production<sup>24</sup>. Our finding that  $7\alpha,25$ -OHC also induces the migration of T cells *in vitro* indicates a possible role for EBI2 in driving T-cell migration and possibly T/B cell interaction *in vivo*. As for B cells, although  $7\alpha,25$ -OHC showed potent activity in inducing the migration of B cells *in vitro*, additional studies showed that it did not affect LPS, or anti-IgM/anti-CD40-induced activation of B cells (Supplementary Fig. 16b, c), similar to the findings reported with EBI2-deficient B cells<sup>4</sup>. However, it remains to be tested whether these ligands affect other aspects of B-cell activation including isotype switching, as it was shown that  $25$ -OHC inhibited IgA production in mice<sup>20</sup>. Because EBI2 is also expressed by dendritic cells, macrophages, NK cells and neutrophils, it is intriguing to speculate that these cells may also be regulated by  $7\alpha,25$ -OHC, in addition to serving as producers of the ligand. To this end, it is interesting to speculate that within the follicle our finding supports the notion that CD169<sup>+</sup> macrophages may be the producers of EBI2 ligands<sup>3</sup>. Extra-follicular areas are rich in NK cells and neutrophils<sup>24,25</sup>. These cells may be the main EBI2 ligand producers and  $7\alpha,25$ -OHC may be the most likely endogenous EBI2 ligand promoting the trafficking of early plasmablasts to extra-follicular areas. Nonetheless, on the basis of our current findings, the observed defect of EBI2-deficient mice in T-cell-dependent antibody production<sup>3,4</sup> might be largely due to a defect in trafficking of a number of immune cells including B cells, T cells and dendritic cells.

We identified  $7\alpha,25$ -OHC as the most likely endogenous ligand for EBI2. Consistent with our observation, a similar finding is reported in

an accompany paper<sup>26</sup>.  $7\alpha,25$ -OHC showed the highest potency at EBI2, with an EC<sub>50</sub> value in the picomolar range. *In vitro* and *in vivo* studies showed that  $7\alpha,25$ -OHC can serve as a chemokine directing migration of B cells, T cells and dendritic cells. *In vivo*,  $7\alpha,25$ -OHC-desensitized B cells, like EBI2-deficient B cells, showed reduced homing to follicular areas of the spleen. In addition, oral administration of clotrimazole, a CYP7B1 inhibitor, reduced  $7\alpha,25$ -OHC levels in the mouse spleen, and promoted adoptively transferred activated B cells to localize primarily at the T/B boundary, which mimics the phenotype of activated EBI2-deficient B cells. Upon *in vivo* injection of LPS, two key enzymes (which are CH25H and CYP7B1) in the biosynthesis of  $7\alpha,25$ -OHC are highly upregulated in spleen dendritic cells, macrophages, and especially in NK cells and neutrophils, indicating that these cells may be the main producers of EBI2 ligands. Our data open up the possibility that oxysterols or their synthetic derivatives may offer therapeutic benefits either as adjuvants or immune modulators for inflammation and autoimmune diseases.

## METHODS SUMMARY

Tissues were extracted using organic solvents followed by normal-phase HPLC. Ligand activity was monitored by GTP- $\gamma$ S binding assay. Structure analysis was done by mass spectrometry and NMR. GTP- $\gamma$ S binding, radioligand binding, cAMP accumulation and receptor internalization studies were used to characterize the receptor. *In vitro* and *in vivo* cell migration assays were used to test the chemotactic effect of  $7\alpha,25$ -OHC on various immune cells.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 8 January; accepted 17 May 2011.**

1. Birkenbach, M., Josefsen, K., Yalamanchili, R., Lenoir, G. & Kieff, E. Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled peptide receptors. *J. Virol.* **67**, 2209–2220 (1993).
2. Pereira, J. P., Kelly, L. M., Xu, Y. & Cyster, J. G. EBI2 mediates B cell segregation between the outer and centre follicle. *Nature* **460**, 1122–1126 (2009).
3. Pereira, J. P., Kelly, L. M. & Cyster, J. G. Finding the right niche: B-cell migration in the early phases of T-dependent antibody responses. *Int. Immunol.* **22**, 413–419 (2010).
4. Gatto, D., Paus, D., Basten, A., Mackay, C. R. & Brink, R. Guidance of B cells by the orphan G protein-coupled receptor EBI2 shapes humoral immune responses. *Immunity* **31**, 259–269 (2009).
5. Goodnow, C. C., Vinuesa, C. G., Randall, K. L., Mackay, F. & Brink, R. Control systems and decision making for antibody production. *Nature Immunol.* **11**, 681–688 (2010).
6. Chan, T. D. *et al.* *In vivo* control of B-cell survival and antigen-specific B-cell responses. *Immunol. Rev.* **237**, 90–103 (2010).
7. Schroeffer, G. J. Jr. Oxysterols: modulators of cholesterol metabolism and other processes. *Physiol. Rev.* **80**, 361–554 (2000).
8. Björkhem, I. & Diczfalusy, U. Oxysterols: friends, foes, or just fellow passengers? *Arterioscler. Thromb. Vasc. Biol.* **22**, 734–742 (2002).
9. Zhao, C. & Dahlman-Wright, K. Liver X receptor in cholesterol metabolism. *J. Endocrinol.* **204**, 233–240 (2010).
10. Willy, P. J. *et al.* LXR, a nuclear receptor that defines a distinct retinoid response pathway. *Genes Dev.* **9**, 1033–1045 (1995).
11. Jin, L. *et al.* Structural basis for hydroxycholesterols as natural ligands of orphan nuclear receptor ROR $\gamma$ . *Mol. Endocrinol.* **24**, 923–929 (2010).
12. Russell, D. W. Oxysterol biosynthetic enzymes. *Biochim. Biophys. Acta* **1529**, 126–135 (2000).
13. Smith, L. L. Cholesterol autooxidation 1981–1986. *Chem. Phys. Lipids* **44**, 87–125 (1987).
14. Rose, K. A. *et al.* Cyp7b, a novel brain cytochrome P450, catalyzes the synthesis of neurosteroids  $7\alpha$ -hydroxy dehydroepiandrosterone and  $7\alpha$ -hydroxy pregnenolone. *Proc. Natl Acad. Sci. USA* **94**, 4925–4930 (1997).
15. Paila, Y. D. & Chattopadhyay, A. Membrane cholesterol in the function and organization of G-protein coupled receptors. *Subcell. Biochem.* **51**, 439–466 (2010).
16. Liu, C. *et al.* Identification of relaxin-3/INSL7 as an endogenous ligand for the orphan G-protein-coupled receptor GPCR135. *J. Biol. Chem.* **278**, 50754–50764 (2003).
17. Lund, E. G., Kerr, T. A., Sakai, J., Li, W. P. & Russell, D. W. cDNA cloning of mouse and human cholesterol 25-hydroxylases, polytopic membrane proteins that synthesize a potent oxysterol regulator of lipid metabolism. *J. Biol. Chem.* **273**, 34316–34327 (1998).
18. Cali, J. J. & Russell, D. W. Characterization of human sterol 27-hydroxylase. A mitochondrial cytochrome P-450 that catalyzes multiple oxidation reaction in bile acid biosynthesis. *J. Biol. Chem.* **266**, 7774–7778 (1991).

19. Diczfalussy, U. *et al.* Marked upregulation of cholesterol 25-hydroxylase expression by lipopolysaccharide. *J. Lipid Res.* **50**, 2258–2264 (2009).
20. Bauman, D. R. *et al.* 25-Hydroxycholesterol secreted by macrophages in response to Toll-like receptor activation suppresses immunoglobulin A production. *Proc. Natl Acad. Sci. USA* **106**, 16764–16769 (2009).
21. Napolitani, G., Rinaldi, A., Bertoni, F., Sallusto, F. & Lanzavecchia, A. Selected Toll-like receptor agonist combinations synergistically trigger a T helper type 1–polarizing program in dendritic cells. *Nature Immunol.* **6**, 769–776 (2005).
22. Reif, K. *et al.* Balanced responsiveness to chemoattractants from adjacent zones determines B-cell position. *Nature* **416**, 94–99 (2002).
23. Kelly, E., Bailey, C. P. & Henderson, G. Agonist-selective mechanisms of GPCR desensitization. *Br. J. Pharmacol.* **153**, S379–S388 (2008).
24. Walzer, T. *et al.* Identification, activation, and selective *in vivo* ablation of mouse NK cells via Nkp46. *Proc. Natl Acad. Sci. USA* **104**, 3384–3389 (2007).
25. Kesteman, N., Vansanten, G., Pajak, B., Goyert, S. M. & Moser, M. Injection of lipopolysaccharide induces the migration of splenic neutrophils to the T cell area of the white pulp: role of CD14 and CXC chemokines. *J. Leukoc. Biol.* **83**, 640–647 (2008).
26. Hannedouche, S. *et al.* Oxysterols direct immune cell migration via EBI2. *Nature*. doi:10.1038/10280 (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank J. Palmer and J. Blevitt for microscopic image acquisition. We thank J. Zhu, R. Luna and S. Nguyen for technical assistance and J. Cowden for scientific discussions.

**Author Contributions** C.L., S.S., J.W., X.V.Y., N.S.M., N.Q., L.K. and T.W.L. designed and conceptualized the research. C.L., S.S., J.W., X.V.Y., N.S.M., C.K., J.Y., S.W.S., N.Q., L.Z. and H.B. performed experiments. C.L., S.S., J.W., X.V.Y., N.S.M., C.K., L.Z., N.Q., L.K. and T.W.L. analysed the data, prepared the figures and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to C.L. ([cliu9@its.jnj.com](mailto:cliu9@its.jnj.com)).

## METHODS

**Materials.** Oxysterols 25-OHC, 7 $\beta$ -OHC, 22(R)-OHC, 24(S)-OHC and 27-OHC were purchased from Sigma-Aldrich. 7 $\alpha$ ,25-OHC and 7 $\beta$ ,25-OHC were synthesized in house. Other oxysterols were purchased from Avanti Polar Lipids. FITC-CD3, APC-B220, APC-CD11b, APC-CD11c, APC-CD49b, APC-Ly6G, FITC-CXCR4, FITC-CXCR5, PE-CCR7, PE-CD69 and PE-CD86 monoclonal antibodies used for FACS were purchased from Becton Dickinson.

**Molecular cloning and recombinant expression of EBI2.** The human (GenBank accession number NM\_004951.4) and mouse (GenBank accession number NM\_183031.2) EBI2 coding regions were PCR-amplified from the human and mouse genomic DNA respectively. Primers P1 (5'-GGTACGGAATTCGCCACCATGGATATACAAATGGCAAACAATTTTACTCC-3') and P2 (5'-ACTAGAGCGGCCGCTCACTTTCCATTTGAAGACTTGGAAATGTATCATCATC-3') were used to amplify the human receptor whereas primers P3 (5'-ACTATCTCTAGAGCCACCATGGCTAACAATTTCACTACCCCACTGGCAA-3') and P4 (5'-ACTCTGGCGGCCGCTTACCTTCCATTGGAGGCCTTGGAGTG-3') were used for the mouse receptor. The resulting DNAs were cloned into a mammalian expression vector (pCIneo, Promega) and the insert sequences confirmed by DNA sequencing (Eton Bioscience). The EBI2 expressing constructs were transiently co-expressed in COS7 cells with a human GO2 expression construct, which contains the human GO2 coding region in a mammalian expression vector pcDNA3.1/zeo (Invitrogen). Two days after transfection, the cells were harvested and cell pellets were kept at -80 °C for later radioligand binding and GTP- $\gamma$ S binding assays. A stable cell line expressing human EBI2 was established by transfecting SK-N-MC/CRE- $\beta$ -gal cells followed by selection under G418.

**GTP- $\gamma$ S binding assays.** These were performed as described previously<sup>16</sup>.

**Converting [<sup>3</sup>H]25-OHC into [<sup>3</sup>H]7 $\alpha$ ,25-OHC.** <sup>3</sup>H-labelled 7 $\alpha$ ,25-OHC was generated from <sup>3</sup>H-labelled 25-OHC (specific activity: 80 Ci mmol<sup>-1</sup>). Briefly, COS7 cells were transiently transfected with a recombinant CYP7B1 (GenBank accession NM\_004820.3) expression plasmid cloned in pcDNA3.1 (Invitrogen). Two days after transfection, the cell culture medium was replaced with serum-free DMEM containing <sup>3</sup>H-labelled 25-OHC and the cells were then cultured overnight. The conditioned medium was collected, dried in a lyophilizer, re-dissolved in methanol and then purified by a reversed-phase HPLC.

**cAMP accumulation assay.** SK-N-MC/CRE- $\beta$ -gal cells stably expressing human EBI2 were seeded in 96-well plates at a density of 50,000 cells per well. Different EBI2 ligands at various concentrations were added to cells to stimulate the receptor. Forskolin was added to a final concentration of 5  $\mu$ M to stimulate cAMP synthesis. The reaction was carried out at room temperature for 20 min and cAMP concentration was measured using cAMP Flash Plate (Perkin Elmer).

**Radioligand binding assay.** COS7 cells transiently expressing EBI2 or control cells without EBI2 expression were seeded in 24-well plates 1 day after transfection. Two days after transfection, the cells were washed with binding buffer (DMEM plus 50 mM HEPES and 0.1% fatty acid free BSA) and then incubated with 500  $\mu$ l binding buffer in the presence of various concentrations of [<sup>3</sup>H]7 $\alpha$ ,25-OHC either with or without unlabelled ligands as competitors. For saturation binding studies, [<sup>3</sup>H]7 $\alpha$ ,25-OHC (specific activity, 80 Ci mmol<sup>-1</sup>) was added to the binding buffer at various concentrations up to 5 nM. Nonspecific binding was assessed by adding unlabelled 7 $\alpha$ ,25-OHC. For competition binding studies, [<sup>3</sup>H]7 $\alpha$ ,25-OHC with 100,000 c.p.m. was added to each well. Different compounds at various concentrations were added to wells as competitors. The bindings were allowed at room temperature for 1 h. The cells were then washed four times with ice-cold PBS. The bound radio-ligand was solubilized by 500  $\mu$ l of 0.1 N NaOH plus 0.2% SDS and counted in a scintillation counter. Each data point was performed in triplets. The results were analysed using Graphpad program (Graphpad).

**Identification and purification of ligand for EBI2.** Initial ligand activity for EBI2 was identified from rat spleen extract. Five grams of rat spleen was extracted with 40 ml of cold ethanol. The ethanol extract was dried in a Buchi Rotovap (New Castle). The dried material was re-dissolved in 20 ml of 0.1% TFA, loaded onto a C18 Cartridge, washed with 50% acetonitrile plus 0.1% TFA, and eluted with 80% acetonitrile plus 0.1% TFA. The eluant was lyophilized, reconstituted in water, diluted at various dilution factors, and then tested for activity using EBI2-expressing membranes in a GTP- $\gamma$ S binding assay. Membrane from cells without EBI2 expression served as the control. To purify and identify the ligand for EBI2, 2 kg of cryofrozen pig spleen purchased from Pel-Freez Biologicals was homogenized with 8 volumes of cold (-30 °C) ethanol. The homogenate was stirred at 4 °C for 2 h and then centrifuged at 15,000g for 1 h. The supernatant was concentrated using a Buchi Rotovap and then dissolved with 500 ml of methanol/chloroform (30:70 v/v). The mixture was allowed to sediment at 4 °C overnight and the clear supernatant was again dried using a Buchi Rotovap. The resulting dried material was divided evenly into two 18  $\times$  150 mm test tubes. To each test tube, 10 ml of water and 10 ml of ethylacetate (EA) were added and the tube was vortexed for 1 min and then allowed to form clear separation between the water and EA layers. The EA layer

was collected, and the water phase was further extracted with 2  $\times$  10 ml EA. The EA phases were combined and dried under nitrogen gas. The above dried material was dissolved in a mixture of 6 ml chloroform, 5.4 ml hexanes, and 0.6 ml isopropanol. The solution was pressed through a Millipore Millex-GV 2.2  $\mu$ m syringe filters and the clear filtered solution was fractionated by normal phase HPLC. Preparative normal phase HPLC runs were carried out on a Shimadzu LC-8A-based system with a Diol column (250  $\times$  30 mm, 5  $\mu$ m particle, 60 Å pore, Princeton Chromatography Inc.). The mobile phase A was 95% hexanes plus 5% isopropanol and the mobile phase B was 85% isopropanol plus 15% methanol. The mobile phase flowed at a rate of 27 ml min<sup>-1</sup> to deliver a 1.6% B per min gradient after 5 min equilibration at 0% B. 3 ml of sample were injected and fractions were collected at 4 tubes per min. 200  $\mu$ l of the each fraction was dried using N<sub>2</sub> gas, re-dissolved in 10  $\mu$ l DMSO and tested in GTP- $\gamma$ S binding assay to stimulate EBI2. The active fractions were pooled and dried by N<sub>2</sub> gas and subjected to GCMS and NMR analysis.

**GCMS chemicals and reagents.** GC-grade pyridine, and GCMS derivatization reagent *N*-methyl-*N*-trimethylsilyltrifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS), were purchased from Sigma-Aldrich.

**Sample derivatization and GCMS analysis.** Either N<sub>2</sub>-dried pig spleen extract after chromatography or standard oxysterol solid (~10  $\mu$ g) was placed in a glass GC vial. To each sample, 10  $\mu$ l of pyridine and 90  $\mu$ l derivatization agent (MSTFA + 1% TMCS) were added. The vial was then vortexed for 30 s and then placed in (floating) a 50 °C water bath. After 1 h, the vial was taken out of the water bath and 1  $\mu$ l sample was injected directly into the GCMS in splitless mode. The GCMS used was an Agilent 7890A GC with 7693 injector and 5975C MSD controlled by Chemstation software (Agilent). The GC column used was an Agilent 122-5532G DB5-MS+DG 30 m  $\times$  0.25 mm, 0.25  $\mu$ m film. GC oven program: 1.2533 ml min<sup>-1</sup> constant flow helium; oven ramp from 60 °C (1 min hold) to 325 °C at 10 °C per min, 10-min hold before cool-down, 37.5 min run time. EI mass spectra were acquired in scan mode from 50 to 600 atomic mass unit (a.m.u.). Data processing was done in Chemstation software, and mass spectral searches were done with NIST spectral library (version 2.0f, build 22 October 2009).

**NMR studies.** Dried samples or standard oxysterols were dissolved in a mixed solvent (0.5 ml CDCl<sub>3</sub> + 0.1 ml CD<sub>3</sub>OD) in 5-mm NMR tubes. Proton NMR spectra were acquired on a Bruker Avance 600 MHz NMR spectrometer with a 5-mm BBO probe using TopSpin 2.0 software. Typical parameters: pulse width = 4  $\mu$ s; pre-pulse delay = 5 s; spectral width = 18 p.p.m.; chemical shift reference = TMS.

**Synthesis of 7 $\alpha$ ,25-OHC.** 7 $\alpha$ ,25-OHC was prepared from commercially available 25-OHC according to the method reported previously<sup>27</sup>. The corresponding 7 $\beta$ ,25-OHC was prepared from the 7-oxo,25-OHC by a stereoselective Luche reduction<sup>28</sup> as detailed in the Supplementary Information. The resulting compounds were confirmed by MS and NMR analysis as described in Supplementary Information.

**Ligand-induced EBI2 internalization.** CHO cells transiently expressing EBI2 with a V5 tag at the N terminus were first incubated with a monoclonal anti-V5 antibody (Invitrogen) at 5  $\mu$ g ml<sup>-1</sup> (diluted in cell culture medium) for 20 min at 37 °C and then stimulated with 1  $\mu$ M 7 $\alpha$ ,25-OHC or a buffer control for 30 min. The cells were then washed with PBS and fixed with 10% formalin. Cy3-labelled goat anti-mouse secondary antibody (1  $\mu$ g ml<sup>-1</sup>) was then used to detect the first antibody. The receptor localization was then visualized under a fluorescent confocal microscope. CHO cells without recombinant EBI2 expression were used as the negative control for staining.

**Mice.** C57BL/6J (B6) mice, approximately 20 g body weight, were purchased from The Jackson Laboratory. The EBI2 knockout (KO) mice on a B6/129 background were ordered from Deltagen, in which the *Ebi2* gene was disrupted by replacing the coding region 45 to 646 (602 bp) with a Neo cassette (7,780 bp) (Supplementary Fig. 13). All studies in mice have been carried out in accordance with the Declaration of Helsinki and/or with the Guide for the Care and Use of Laboratory Animals as adopted and promulgated by the US National Institutes of Health.

**Purification of various types of mouse splenocyte.** B6 mice were treated with saline or LPS, 5 mg kg<sup>-1</sup>, intraperitoneally. Two hours after treatment, the spleens were harvested and treated with collagenase before preparation of cell suspension. Various subtypes of splenocyte were then prepared by positive selection using Miltenyi Biotec magnetic beads for anti-CD3 (pan-T cells), anti-B220 (B cells), anti-CD11b (macrophages), anti-CD11c (DC), anti-CD49b (NK cells) and anti-Ly6G (Neutrophils). The purity of isolated cells was examined by FACS using FITC-CD3, APC-B220, APC-CD11b, APC-CD11c, APC-CD49b, APC-Ly6G and found to be at least 95% pure.

**Quantitative analysis of mRNA expression in different splenocytes.** B6 mice were treated with saline or LPS (5 mg kg<sup>-1</sup>, intraperitoneally) for 2 h, spleens were isolated, subtype of splenocytes purified as described above, and total RNA was extracted. RNAs were reverse-transcribed into cDNAs using random primer by standard protocols. cDNAs for mouse *Ebi2*, *Cyp7b1*, *Ch25h*, *Cyp27a* and  $\beta$ -actin were then PCR amplified using gene-specific primers for *Ebi2* (forward,



5'-GACATCCTGTTTACCACAGCT-3'; reverse, 5'-AGACCAGAATCCAGACGGACA-3'), *Cyp7b1* (forward, 5'-TCAGGAAAGGCAAGATCTGCTGA-3'; reverse, 5'-CCTGTTGACTGCAGGAACTGTCA-3'), *Ch25h* (forward, 5'-CTGCCTGCTGCTCTTCGACA-3'; reverse, 5'-CCGACAGCCAGATGTTAATCA-3'), *Cyp27a1* (forward, 5'-CCCTTTTGAAGCGATACCTG-3'; reverse, 5'-GTCAGTGTGTTGGATGTCTGT-3') and  $\beta$ -actin (forward, 5'-ACAACGGCTCCGGCATGTGCA-3'; reverse, 5'-GTGTGGTGCCAGATCTTCTCA-3'), respectively. Purified PCR products for each gene were diluted at various concentrations as templates for amplification and serve the standards for quantification. qPCR amplification and quantification were performed using a TaqMan PCR machine (Applied Biosystems).

**Extraction and quantification of 7 $\alpha$ ,25-OHC from mouse spleens.** For the LPS treatment group, B6 mice were treated with saline or LPS, 5 mg kg<sup>-1</sup> for 8 h, and spleens were isolated. For clotrimazole treatment groups, B6 mice were treated with vehicle (20% hydroxypropyl  $\beta$ -cyclodextrin) or clotrimazole (100 mg kg<sup>-1</sup>, dosed twice orally, 8 h apart). Sixteen hours after the second treatment, the spleens were isolated and frozen in liquid nitrogen immediately. Each mouse spleen was weighed and then homogenized in 3 ml of 1 N KOH in methanol plus 2 mM butylated hydroxytoluene (for preventing cholesterol auto-oxidation) and 2 mM EDTA. The extract was incubated at 37 °C for 1 h. 3 ml of NaCl (100 mM) was then added to the mixture and the solution was neutralized using concentrated phosphoric acid. Six millilitres of chloroform was then added to solution, extracted for 3 min and centrifuged at 3,000g for 10 min. The chloroform phase was collected. The aqueous phase was extracted one more time using 6 ml of chloroform/methanol (2:1) and the chloroform phase was collected. The chloroform phases were combined and dried using a speed vacuum drier. The dried extracts were subjected to normal phase HPLC fractionation with the following conditions. The dried extract was dissolved in 0.35 ml of a mixed solvent (1:1 v/v chloroform, methanol, mobile phase A). The solution was filtered through a Millipore PVDF 0.22  $\mu$ m spin filter and then fractionated by normal phase HPLC under the following conditions: Waters Alliance 2690 HPLC; Princeton Chromatography Inc. Diol column (250  $\times$  10 mm, 5  $\mu$ m particle, 60 Å pore); mobile phase A: 95% hexanes and 5% isopropanol; mobile phase B: 85% isopropanol, 15% methanol, 0.2% acetic acid, 0.05% triethylamine; flow rate 3 ml min<sup>-1</sup>; gradient: 0–5 min 0% B, 5–35 min 0–50% B; injection volume: 0.3 ml; fraction collection: 4 tubes per min. The collected fractions were dried in Genevac-HT, dissolved in 100  $\mu$ l of 50% DMSO and tested for ligand activity against EBI2 expressing membrane. 7 $\alpha$ ,25-OHC was quantified using EBI2-expressing membrane in a GTP- $\gamma$ S binding assay with known concentrations of 7 $\alpha$ ,25-OHC as the standards for quantification. The remaining active fractions from each group after activity tests were pooled, dried by lyophilization and subjected to GCMS analysis.

**Sample derivatization and GCMS analysis.** To each dried sample, 25 ng coprostanol was included as reference. Chemical derivatization was initiated by adding 100  $\mu$ l of the reaction mix (MSTFA 1 ml, NH<sub>4</sub>I 4 mg, dithioerythritol 4 mg, dichloromethane 0.1 ml) to the sample. After 30 min at 60 °C in a water bath, 4  $\mu$ l of the derivatized sample was subjected to GCMS analysis<sup>29</sup>. The GC oven temperature ramp started from 60 °C (1-min hold) to 280 °C at 10 °C min<sup>-1</sup>, from

280 °C to 325 °C at 5 °C min<sup>-1</sup>, 10-min hold before cool-down, 42 min total run time. EI mass spectra were acquired in SIM mode for the following ions: 73.1, 131.1, 370.4, 456.4, 544.5, 546.4.

**In vitro migration assay.** Red-blood-cell-lysed mouse splenocyte suspension in RPMI plus 10% fetal calf serum (FCS) were prepared from spleens of *Ebi2*<sup>+/-</sup> (B6/129), *Ebi2*<sup>-/-</sup> (B6/129) and wild-type littermate (B6/129) mice, then activated with 1  $\mu$ g ml<sup>-1</sup> LPS (055:B5, Sigma) plus 1  $\mu$ g ml<sup>-1</sup> LPS (0111:B4, Sigma) or 1  $\mu$ g ml<sup>-1</sup> anti-CD3 plus 5  $\mu$ g ml<sup>-1</sup> anti-CD28 overnight. Activated splenocytes were washed once and re-suspended at  $\sim 1 \times 10^7$  cells ml<sup>-1</sup> in RPMI with 10% FCS, 100  $\mu$ l of the suspension was added to a top chamber of the 24-well transwell plate (Corning, 3421), whereas the bottom chambers contained 600  $\mu$ l of various concentrations of 7 $\alpha$ ,25-OHC. Migration was allowed for 2 h at 37 °C and the migrated cells were then counted and analysed using FACS after staining for B220, CD4, CD8, CD11c, CD11b, CD49b, or Ly6G. Nonspecific migration (cells migration to medium only) was subtracted to yield specific migration towards compound.

For 7 $\alpha$ ,25-OHC pre-treatment *in vitro*, B220<sup>+</sup> cells were isolated from B6 splenocytes by positive magnetic separation using an AutoMacs separator (Miltenyi biotech) and activated with 1  $\mu$ g ml<sup>-1</sup> anti-IgM (goat anti-mouse IgM, Sigma) plus 5  $\mu$ g ml<sup>-1</sup> anti-CD40 (Biolegend, clone 1C10) overnight. Activated B cells were then washed once and treated with various concentrations of 7 $\alpha$ ,25-OHC as indicated or DMSO overnight. On day 3, cells were washed three times and migration towards 10 nM 7 $\alpha$ ,25-OHC, 200 ng ml<sup>-1</sup> CXCL12, 200 ng ml<sup>-1</sup> CXCL13, or 100 ng ml<sup>-1</sup> CCL21 were tested as described above.

**In vivo migration of adoptively transferred B cells.** B220<sup>+</sup> spleen cells were isolated and activated with 1  $\mu$ g ml<sup>-1</sup> anti-IgM (goat anti mouse IgM, Sigma) plus 5  $\mu$ g ml<sup>-1</sup> anti-CD40 (Biolegend, clone 1C10) overnight as described above. Activated B cells were labelled with CFSE (Invitrogen), re-suspended in DPBS + 0.1% BSA at  $6 \times 10^6$  cells ml<sup>-1</sup>, and adoptively transferred (200  $\mu$ l, intravenous) into B6 recipient mice. For desensitization study, before transfer, cells were treated with DMSO or 1  $\mu$ M 7 $\alpha$ ,25-OHC for 1.5 h. Spleens and blood samples were harvested at indicated time points. Transferred B cells and follicles in spleen cryosections were detected by IHC using anti-fluorescein-AP (Roche) and biotin-anti-IgD/biotin-anti-B220 (Becton Dickinson) and Streptavidin-HRP. Transferred B cells in blood samples were analysed by FACS analysis. The total numbers of CFSE<sup>+</sup> B cells found in representative follicles ( $n = 9$ –12) as well as in B zone or T/B boundary (including T-zone) were counted (shown in Supplementary Fig. 17) and the percentages in specific follicular areas calculated.

27. Li, D. & Spencer, T. A. Synthesis of 7 $\alpha$ -hydroxy derivatives of regulatory sterols. *Steroids* **65**, 529–535 (2000).
28. Št'astrná, E. *et al.* Stereoselectivity of sodium borohydride reduction of saturated steroidal ketones using conditions of Luche reduction. *Steroids* **75**, 721–725 (2010).
29. Kumar, B. S. *et al.* Gas chromatography-mass spectrometry-based simultaneous quantitative analytical method for urinary oxysterols and bile acids in rats. *Anal. Biochem.* **408**, 242–252 (2011).

# Oxysterols direct immune cell migration via EBI2

Sébastien Hannedouche<sup>1\*</sup>, Juan Zhang<sup>2\*</sup>, Tangsheng Yi<sup>3\*</sup>, Weijun Shen<sup>4,5</sup>, Deborah Nguyen<sup>4</sup>, João P. Pereira<sup>3</sup>, Danilo Guerini<sup>6</sup>, Birgit U. Baumgarten<sup>7</sup>, Silvio Roggo<sup>8</sup>, Ben Wen<sup>4</sup>, Richard Knochenmuss<sup>2</sup>, Sophie Noël<sup>1</sup>, Francois Gessier<sup>8</sup>, Lisa M. Kelly<sup>3</sup>, Mirka Vanek<sup>7</sup>, Stephane Laurent<sup>7</sup>, Inga Preuss<sup>7</sup>, Charlotte Miault<sup>8</sup>, Isabelle Christen<sup>2</sup>, Ratna Karuna<sup>2</sup>, Wei Li<sup>4</sup>, Dong-In Koo<sup>5</sup>, Thomas Suply<sup>7</sup>, Christian Schmedt<sup>4</sup>, Eric C. Peters<sup>4</sup>, Rocco Falchetto<sup>2</sup>, Andreas Katopodis<sup>6</sup>, Carsten Spanka<sup>8</sup>, Marie-Odile Roy<sup>1</sup>, Michel Detheux<sup>1</sup>, Yu Alice Chen<sup>4</sup>, Peter G. Schultz<sup>4</sup>, Charles Y. Cho<sup>4</sup>, Klaus Seuwen<sup>7</sup>, Jason G. Cyster<sup>3</sup> & Andreas W. Sailer<sup>7</sup>

**Epstein–Barr virus-induced gene 2 (EBI2, also known as GPR183) is a G-protein-coupled receptor that is required for humoral immune responses; polymorphisms in the receptor have been associated with inflammatory autoimmune diseases<sup>1–3</sup>. The natural ligand for EBI2 has been unknown. Here we describe the identification of 7 $\alpha$ ,25-dihydroxycholesterol (also called 7 $\alpha$ ,25-OHC or 5-cholesten-3 $\beta$ ,7 $\alpha$ ,25-triol) as a potent and selective agonist of EBI2. Functional activation of human EBI2 by 7 $\alpha$ ,25-OHC and closely related oxysterols was verified by monitoring second messenger readouts and saturable, high-affinity radioligand binding. Furthermore, we find that 7 $\alpha$ ,25-OHC and closely related oxysterols act as chemoattractants for immune cells expressing EBI2 by directing cell migration *in vitro* and *in vivo*. A critical enzyme required for the generation of 7 $\alpha$ ,25-OHC is cholesterol 25-hydroxylase (CH25H)<sup>4</sup>. Similar to EBI2 receptor knockout mice, mice deficient in CH25H fail to position activated B cells within the spleen to the outer follicle and mount a reduced plasma cell response after an immune challenge. This demonstrates that CH25H generates EBI2 biological activity *in vivo* and indicates that the EBI2–oxysterol signalling pathway has an important role in the adaptive immune response.**

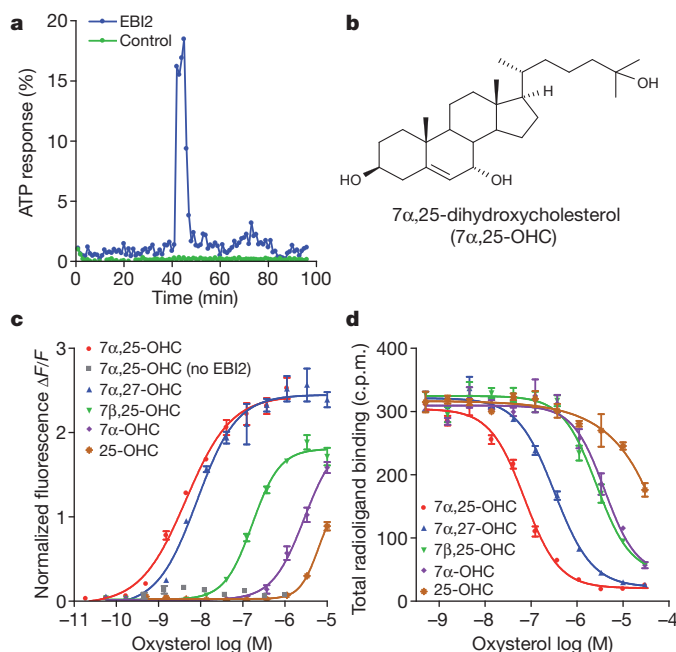
Screening of a liver extract from a peritonitis sheep sepsis model in a functional assay format detecting release of intracellular calcium identified a bioactivity specific for human EBI2 (Fig. 1a). After characterization of biophysical properties (for example, protease resistance), 2 kg of sheep liver were extracted and fractionated over six steps of liquid chromatography (Supplementary Fig. 1), which yielded pure fractions potentially activating the human EBI2 receptor.

Mass spectrometry (MS) analysis of the active fractions using nano electrospray ionization–Fourier transform MS (NanoESI–FTMS) showed four peaks with significant intensities in both step 5 (Supplementary Fig. 1e, Active 1) and step 6 (Supplementary Fig. 1f, Active 2), which were negligible in control samples (Supplementary Fig. 2). Ions with masses  $m/z$  383.3303, 401.3408 and 441.3332 were observed in positive ion mode, whereas  $m/z$  477.3584 was in negative ion mode. The ions corresponding to masses  $m/z$  401.3408 and 383.3303 could be assigned to a protonated molecule with loss of one and two water molecules— $[M + H - 2H_2O]^+$  and  $[M + H - H_2O]^+$ —whereas masses  $m/z$  441.3332 and 477.3582 could be assigned to its sodium and acetate adducts, respectively. Accurate mass measurements allowed assigning a  $C_{27}H_{46}O_3$  molecular formula to the identified 418.3447 molecular weight.

A database query with  $C_{27}H_{46}O_3$  or mass 418.3447 provided mainly oxidized cholesterol or vitamin D derivatives. We found around 50 candidates when querying public databases (CHEBI, HMDB, Kegg, LipidMap, Metline, BioCyc and LipidBank) but very few were

commercially available. Among the tested compounds two related oxysterols, cholest-5-ene-3 $\beta$ ,7 $\beta$ ,25-triol (7 $\beta$ ,25-OHC) and cholest-5-ene-3 $\beta$ ,7 $\beta$ -diol (7 $\beta$ -OHC), induced specific activation of the human EBI2 receptor with half-maximum effective concentration ( $EC_{50}$ ) values close to 50 nM and 15  $\mu$ M, respectively, whereas the 7-methoxy derivate of 7 $\beta$ -OHC did not activate EBI2, indicating some structure activity relationship (Supplementary Fig. 3).

In contrast to 7 $\alpha$ ,25-OHC, for which a clear metabolic pathway has been described<sup>4</sup>, the generation of 7 $\beta$ ,25-OHC *in vivo* is still unclear<sup>5</sup>.



**Figure 1 | Identification, structure and pharmacological characterization of 7 $\alpha$ ,25-OHC.** **a**, Bioactivity profile measured by mobilization of intracellular calcium of fractions from septic sheep liver extracts tested on EBI2-expressing cells (blue) or tested on a control cell line (green) expressing an unrelated G-protein-coupled receptor called CCRL2. Activity is plotted relative to an ATP response of an endogenous ATP receptor. **b**, Chemical structure of 7 $\alpha$ ,25-dihydroxycholesterol (7 $\alpha$ ,25-OHC). **c**, Calcium mobilization: dose–response curve induced by several related oxysterols in a CHO cell line stably expressing EBI2. Data for 7 $\alpha$ ,25-OHC on the parental cell line lacking expression of EBI2 are given as 7 $\alpha$ ,25-OHC (no EBI2). Data shown are from a duplicate experiment (mean  $\pm$  s.e.). **d**, Radioligand binding assay: displacement of 10 nM [<sup>3</sup>H]7 $\alpha$ ,25-OHC bound to membranes from EBI2-expressing CHO cells by increasing concentrations of unlabelled oxysterols. Data from quadruplicate experiments (mean  $\pm$  s.e.) are shown. c.p.m., counts per min.

<sup>1</sup>Euroscreen S.A., 6041 Gosselies, Belgium. <sup>2</sup>Analytical Sciences, Novartis Institutes for BioMedical Research, 4056 Basel, Switzerland. <sup>3</sup>Howard Hughes Medical Institute and Department of Microbiology and Immunology, University of California San Francisco, California 94143-0414, USA. <sup>4</sup>Genomics Institute of the Novartis Research Foundation, San Diego, California 92121, USA. <sup>5</sup>Department of Chemistry, The Scripps Research Institute, La Jolla, California 92037, USA. <sup>6</sup>Autoimmunity, Transplantation and Inflammation, Novartis Institutes for BioMedical Research, 4056 Basel, Switzerland. <sup>7</sup>Developmental and Molecular Pathways, Novartis Institutes for BioMedical Research, 4056 Basel, Switzerland. <sup>8</sup>Global Discovery Chemistry, Novartis Institutes for BioMedical Research, 4056 Basel, Switzerland.

\*These authors contributed equally to this work.

**Table 1 | Pharmacological characterization of different oxysterols**

Compound	Ca <sup>2+</sup> fluorescence, EC <sub>50</sub> (nM)	GTP-γ <sup>35</sup> S binding, EC <sub>50</sub> (nM)	[ <sup>3</sup> H]7α,25-OHC binding, IC <sub>50</sub> (nM)
7α,25-OHC	2 ± 2	8 ± 4	70 ± 4
7α,27-OHC	5 ± 4	5 ± 2	362 ± 32
7β,25-OHC	121 ± 54	221 ± 105	2,632 ± 110
7β,27-OHC	310 ± 261	125 ± 66	>10,000
7α-OHC	1,453 ± 1,163	498 ± 267	3,536 ± 431
25-OHC	3,032 ± 2,645	727 ± 232	>10,000
7β-OHC	>10,000	>10,000	>10,000
Cholesterol	>10,000	>10,000	>10,000

Pharmacological characterization was carried out via calcium fluorescence (FLIPR), GTP-γ<sup>35</sup>S binding and radioligand binding assay. *n* = 3–4; values are given with standard deviation of the mean. EC<sub>50</sub>, half-maximum effective concentration; IC<sub>50</sub>, half-maximum inhibitory concentration.

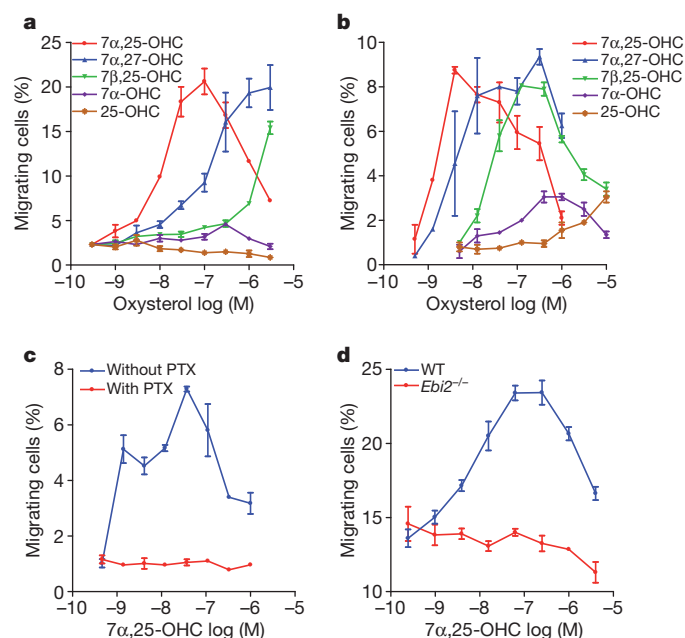
This information prompted us to synthesize 7α,25-OHC (Fig. 1b), which proved to be more potent than 7β,25-OHC (Fig. 1c). While the purification from septic sheep liver was in progress, we also initiated a parallel approach using pig liver tissue (starting material 10 kg). After five liquid chromatography separation steps, ions with masses *m/z* 365.3202, 383.3311, 401.3411 and 441.333 were observed in positive ion mode, rationalized as M + H – 3H<sub>2</sub>O, M + H – 2H<sub>2</sub>O, M + H – H<sub>2</sub>O and M + Na, respectively. To enable <sup>1</sup>H-NMR analysis of the active fraction, a scale-up of the purification using 35 kg pig liver tissue yielded about 100 μg of substantially pure EBI2 ligand. <sup>1</sup>H-NMR analysis comparing EBI2 ligand purified from pig liver with synthetic 7α,25-OHC confirmed the molecular identity of the two compounds (Supplementary Fig. 4), indicating that 7α,25-OHC is the naturally occurring receptor ligand.

After identification of 7α,25-OHC as an EBI2 agonist, we conducted a pharmacological characterization of 7α,25-OHC and closely related oxysterols using two functional second messenger readouts (release of intracellular calcium (Fig. 1c) and binding of GTP-γS; see Table 1). Other than the hydroxylation at the 3 position, hydroxylations at the 7 and at the 25 positions were necessary for potent activation of EBI2. At the 7 position the axial α position was much preferred over the equatorial β position. Hydroxylation of the side chain at the 25 position was preferred but hydroxylation at the 27 position also led to a strong activation. Parental cells lacking the EBI2 receptor did not respond to the tested oxysterols (data for 7α,25-OHC are shown in Fig. 1c; data for the other oxysterols are omitted for clarity). To characterize binding of oxysterols to the EBI2 receptor we developed a suitable radioligand. Reduction of 7-keto,25-OHC using tritiated L-selectride led to the formation of [<sup>3</sup>H]7,25-OHC. Separation of the two stereoisomers yielded radio-labelled [<sup>3</sup>H]7α,25-OHC as well as [<sup>3</sup>H]7β,25-OHC (Supplementary Fig. 5a). Using membranes from a cell line stably expressing human EBI2 we found saturable, high-affinity binding of [<sup>3</sup>H]7α,25-OHC to EBI2 (Supplementary Fig. 5b) with a dissociation constant (*K<sub>d</sub>*) of 25 ± 10 nM (*n* = 3). This binding was displaced by increasing concentrations of 7α,25-OHC or closely related oxysterols (Fig. 1d and Table 1). The rank order of the different oxysterol compounds was identical to the functional assays. No specific binding to EBI2 membranes was detected using the other stereoisomer, [<sup>3</sup>H]-7β,25-OHC. Membranes from parental cells lacking EBI2 receptor did not bind [<sup>3</sup>H]7α,25-OHC (data not shown).

Oxysterols have pleiotropic physiological activities: they regulate the expression of genes that participate in both sterol and fat metabolism, serve as substrates for the synthesis of bile acids, and are intermediates in the transfer of sterols from the periphery to the liver<sup>6</sup>. In addition, oxysterol-dependent activation of the nuclear hormone receptor LXR has recently been implicated in the acquired immune response<sup>7</sup>. To assess possible interaction of the most potent EBI2 ligands with other receptors we tested 7α,25-OHC and 7α,27-OHC in a panel of reporter gene assays against eight different nuclear hormone receptors including LXR and thirty-one different G-protein-coupled receptors. In none of these assays did we see a significant interaction between 7α,25-OHC or 7α,27-OHC and the tested partner (see Supplementary Information).

B cells in different stages of maturation show a distinct profile of EBI2 expression. It is abundant in naive B cells and further upregulated

early after their activation by antigen, whereas B cells from germinal centres sharply downregulate EBI2 expression. Once affinity maturation has been completed and the B cells differentiate into plasma or memory cells, EBI2 is again expressed. This expression pattern combined with the recent finding that EBI2 is a key regulator positioning B cells in lymphoid organs<sup>1,2</sup> prompted us to investigate oxysterols as chemoattractants in a transwell assay system. For this experiment we used either an Epstein–Barr-virus (EBV)-infected human B cell line or a Burkitt's lymphoma pre-B-cell line called RS11846 (refs 8, 9). Both cell lines displayed high endogenous EBI2 expression (Supplementary Fig. 6). We found that 7α,25-OHC can potentially attract EBV-infected B cells or RS11846 cells (Fig. 2a, b). Similar to classical chemokines, we found a bell-shaped activity curve in the transwell assay that indicates reduced attraction at high compound concentrations. Oxysterols closely related to 7α,25-OHC can also attract RS11846 cells albeit with lower potency. The rank order of different oxysterols in the migration assay was identical to the rank order in the functional and binding assays. The attraction of EBI2-expressing RS11846 cells could be blocked by pertussis toxin, a reagent that blocks signalling of Gα<sub>i</sub>-coupled receptors (Fig. 2c). Oxysterol-mediated chemoattraction is not limited to B-lineage cells but can also be observed with other immune cells (for example, dendritic cells) expressing EBI2. Messenger RNA expression



**Figure 2 | Oxysterol-mediated immune cell migration.** **a**, Migration of EBV-infected B cells towards 7α,25-OHC and closely related oxysterols. **b**, Migration of RS11846 cells towards 7α,25-OHC and closely related oxysterols. **c**, Oxysterol-mediated RS11846 cell migration in the presence (red) or absence (blue) of pertussis toxin (PTX). **d**, Migration of bone-marrow-derived dendritic cells from wild-type (WT; blue) or *Ebi2*<sup>-/-</sup> mice (red). Data for all four experiments were conducted in a transwell assay system with duplicate (a–c) or quadruplicate (d) samples.

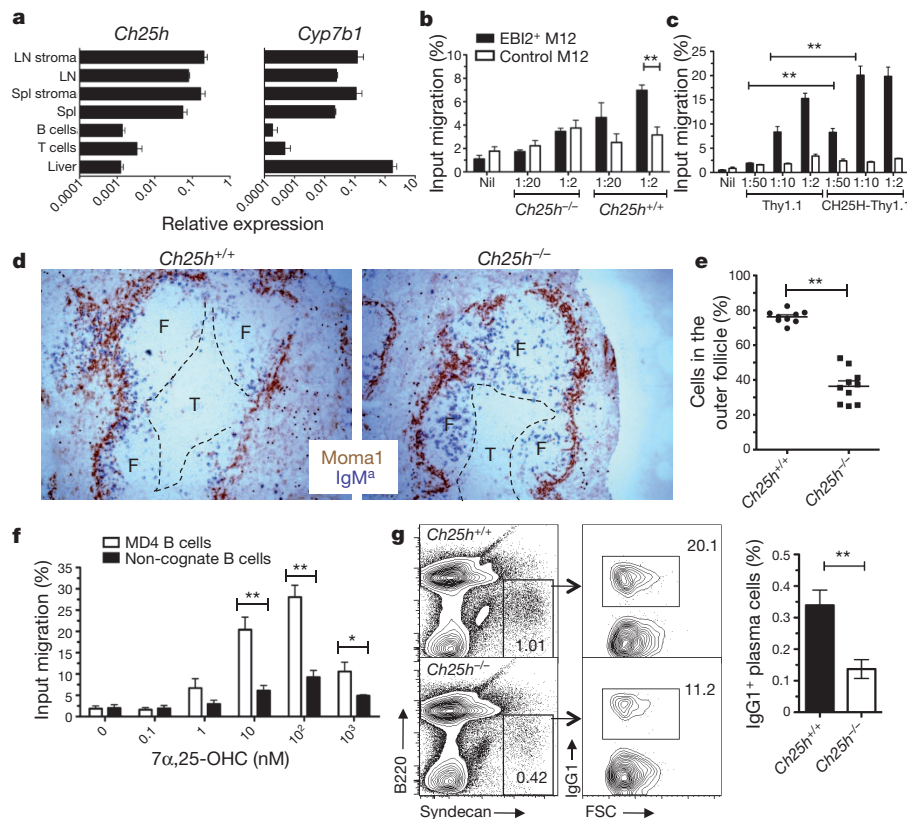


analysis of different human immune cells indicated that various myeloid cells highly express *EBI2* (ref. 10 and Supplementary Fig. 7). Testing of bone-marrow-derived dendritic cells<sup>11</sup> from wild-type and *Ebi2*<sup>-/-</sup> mice in the transwell migration assay showed that only wild-type cells migrate towards 7 $\alpha$ ,25-OHC (Fig. 2d).

Identification of the chemoattractant properties of oxysterols on B cells and the previous findings regarding *EBI2*-deficient B cells<sup>1,2</sup> suggested that ablation of enzymes necessary for ligand production will have an impact on the B-cell trafficking after an immune challenge. The defined biosynthetic pathway of the more potent 7 $\alpha$ ,25-OHC requires the activity of cholesterol 25-hydroxylase (CH25H)<sup>4</sup> and oxysterol 7 $\alpha$ -hydroxylase (CYP7B1)<sup>4</sup>. Both of these enzymes are abundantly expressed in spleen and lymph nodes (Fig. 3a). To determine whether CH25H deficiency affected *EBI2* ligand production, tissue extracts were prepared from wild-type and *Ch25h*<sup>-/-</sup> (ref. 12) spleen and lymph nodes and tested for their ability to attract an *EBI2*-expressing cell line in a bioassay that has subnanomolar sensitivity for 7 $\alpha$ ,25-OHC (Supplementary Fig. 8c). Whereas wild-type tissue extracts showed readily identifiable activity, extracts from *Ch25h*<sup>-/-</sup> mice lacked activity in this bioassay (Fig. 3b). In addition, spleen extracts from mice

reconstituted with bone marrow transduced with a CH25H-expressing retroviral vector showed a greatly increased tendency to attract *EBI2*-expressing cells (Fig. 3c). Consistent with evidence that CH25H expression is increased by lipopolysaccharide (LPS) exposure<sup>13</sup>, the bioactivity in spleen extracts from 6- and 16-h treated mice was increased several fold (Supplementary Fig. 8a, b). Using mass spectrometry we assessed the production of 7 $\alpha$ ,25-OHC in the spleen of unchallenged and LPS-challenged mice (Supplementary Fig. 8d). A peak corresponding to 7 $\alpha$ ,25-OHC was clearly detectable in tissue prepared from LPS-challenged wild-type mice, but was close to the detection limit or undetectable in unchallenged wild-type or *Ch25h*<sup>-/-</sup> mice. The origin of the weak signal detected in some samples derived from *Ch25h*<sup>-/-</sup> mice is not yet clear and could be due to the production of low amounts of 7 $\alpha$ ,25-OHC via enzymes other than CH25H (refs 14–16). These results indicate that in an LPS immune challenge, 7 $\alpha$ ,25-OHC is generated in the spleen in a CH25H-dependent manner.

To test whether 7 $\alpha$ ,25-OHC regulates B-cell migration *in vivo*, we tested *Ch25h*<sup>-/-</sup> mice<sup>12</sup> in an adoptive transfer experiment. *Ch25h*<sup>-/-</sup> or wild-type mice received transfers of hen egg lysozyme (HEL)-specific IgM<sup>a</sup> MD4 B cells and ovalbumin (OVA)-specific OTII T cells



**Figure 3 | CH25H expression regulates *EBI2* bioactivity *in vivo* and is required for mounting a T-cell-dependent antibody response.**

**a**, Quantitative PCR analysis of *Ch25h* and *Cyp7b1* transcript abundance in the indicated tissues or cell preparations, relative to *Hprt* (mean from triplicate samples  $\pm$  s.e.). LN, lymph node; Spl, spleen. **b**, Bioassay of spleen extracts from *Ch25h*<sup>-/-</sup> or wild-type animals using a reporter cell line (M12) with or without *EBI2* (mean from 3 mice  $\pm$  s.e., two experiments). **c**, Bioassay of spleen extracts from mice reconstituted with bone marrow transduced with a CH25H-expressing retroviral vector (CH25H-Thy1.1) or a control vector (Thy1.1) (mean from 3 mice  $\pm$  s.e., representative of 2 similar experiments). **d**, Spleen sections from the indicated mice after immune challenge (day 2). Visualization of antigen-specific transferred (IgM<sup>a</sup>, blue) B cells in *Ch25h*<sup>+/+</sup> mice indicates that many of the activated B cells have moved to the back of the follicle and interfollicular regions, near Moma1<sup>+</sup> marginal metallophilic macrophages (brown). In the *Ch25h*<sup>-/-</sup> mice, the activated B cells largely fail to move to these regions and instead remain near the follicle–T-cell zone interface or move into the follicle. Dashed lines indicate the boundary between follicle (F) and T zone

(T), identified in serial sections using IgD staining. These images are representative of multiple sections from three mice of each genotype. **e**, Fraction of day 2 activated B cells located in the outer follicle. Sections were stained as in **d** and IgM<sup>a</sup> cells within  $\sim 70$   $\mu$ m of the Moma1<sup>+</sup> cells were enumerated and divided by the total number of IgM<sup>a</sup> cells in the follicle. Each point corresponds to an individual follicle. Enumeration was performed on sections from three mice. **f**, Migration of antigen-specific (MD4) B cells and non-cognate B cells from spleens of day 2 immunized wild-type mice in response to 7 $\alpha$ ,25-OHC (mean from 4 mice  $\pm$  s.e., combined from 2 experiments). **g**, Plasma cell response in *Ch25h*<sup>+/+</sup> and *Ch25h*<sup>-/-</sup> mice at day 5 after immunization with sheep red blood cells. Left profiles show spleen cells stained for the indicated markers with plasma cells identified as B220<sup>lo</sup>Syndecan<sup>hi</sup> and further stained to detect intracellular IgG1. Numbers indicate the percentage of total cells in the indicated gate. The right bar graph shows a summary of the data for nine mice of each type (mean  $\pm$  s.e.). For all graphs: \**P* < 0.05, \*\**P* < 0.01 (unpaired student's *t*-test).

at day -1 and were immunized with HEL-OVA at day 0 and analysed at day 2 for their immune response. Visualization of activated B cells (IgM<sup>a</sup>) in the spleen relative to the position of marginal metallophilic macrophages (Moma1) (Fig. 3d) or endogenous follicular B cells (Supplementary Fig. 9) showed that at day 2 after immune challenge in wild-type recipients, many of the activated B cells had moved to the back of the follicle and interfollicular regions whereas in *Ch25h*<sup>-/-</sup> mice activated B cells largely failed to move to these regions and instead remained near the follicle-T-zone interface or moved into the follicle (Fig. 3d, e and Supplementary Fig. 9). The defective distribution of activated B cells in CH25H-deficient mice was similar to that reported for day 2 activated EBI2-deficient B cells<sup>1</sup>. Migration assays with spleen cells prepared at the same time point showed that the day 2 activated wild-type MD4 B cells responded more vigorously than naive endogenous B cells to 7 $\alpha$ ,25-OHC (Fig. 3f). Both the baseline naive B-cell response and the increased response of activated B cells required EBI2 expression (Supplementary Fig. 10) and MD4 B cells that had not been exposed to antigen did not show an augmented response (Supplementary Fig. 10). EBI2 bioactivity was similar in spleen extracts from day 2 immunized and unimmunized mice (Supplementary Fig. 10). *Ch25h* and *Cyp7b1* transcript abundance was low in B and T lymphocytes but enriched in stromal cells (Fig. 3a), indicating that the enzymes are expressed in adhesive myeloid or mesenchymal cells, consistent with studies on other tissues<sup>13,17,18</sup>. Although it has not been possible to determine enzyme distribution in tissue sections with existing antibody reagents and no tools are available to visualize ligand distribution, based on the EBI2-dependent homing properties of activated B cells we speculate that ligand abundance is higher in outer and interfollicular regions compared to the follicle centre. Finally, we tested whether ligand deficiency led to a defect in the plasma cell response. Five days after immunization with a T-dependent antigen (sheep red blood cells), the CH25H-deficient mice showed a ~3-fold defect in the magnitude of their IgG1 plasma cell response (Fig. 3g), an impairment very similar to that observed in EBI2-deficient mice<sup>1,2</sup>.

We have identified the natural ligand for EBI2 and delineated its physiological function as a chemoattractant in the immune system<sup>1,2</sup>. Our results provide an unanticipated link between recent results showing a crucial role for EBI2 in promoting B-cell localization within lymphoid organs, of EBI2 connecting through IRF7 to regulate a macrophage-associated autoimmune disease<sup>3</sup> and studies highlighting how oxysterols shape the innate and adaptive immune responses<sup>12,13,19</sup>. Together with these published reports, our results establish a new physiological dimension for oxysterols as bioactive signalling molecules.

**Note added in proof:** An independent study has obtained similar results identifying oxysterols as EBI2 ligands directing B-cell migration<sup>20</sup>.

## METHODS SUMMARY

**Compounds.** 7 $\alpha$ ,25-OHC (CAS no. 64907-22-8) and 7 $\beta$ ,25-OHC (CAS no. 64907-21-7) were synthesized either in a one-step reaction from the 7-keto precursor (see Supplementary Information), or in a four-step synthesis according to published procedures<sup>21</sup>. 7 $\alpha$ ,27-OHC (CAS no. 144300-24-3) and 7 $\beta$ ,27-OHC (CAS no. 240129-43-5) were purchased from Avanti Polar Lipids Inc. 7 $\alpha$ -OHC (CAS no. 566-26-7) was batch synthesized at Novartis. 7 $\beta$ -OHC (CAS no. 566-27-8), 25-OHC (CAS no. 2140-46-7) and cholesterol (CAS no. 57-88-5) were purchased from Sigma-Aldrich.

**EBI2 recombinant cell line.** The human *EBI2* gene was cloned using RT-PCR from HL-60 cDNA and corresponded to the GenBank accession number NP\_004942. Human *EBI2* was inserted into the pEFIN3 vector at the EcoRI and XbaI sites in the multiple cloning site. Plasmids expressing human EBI2 were stably transfected in CHO-K1 expressing the apo-aequorin and Gq15 to yield CHO/Aeq/Gq15/hEBI2.

**Tissue extract preparation.** Extraction of sheep tissue was done by methanol/water/acetic acid (90:9:1) whereas pig tissue was extracted using hexane/isopropanol (1:1). Methods for separation of crude extracts as well as details on the pharmacological characterization and spectroscopic methods used can be found in Supplementary Information.

Received 22 December 2010; accepted 9 June 2011.

- Pereira, J. P., Kelly, L. M., Xu, Y. & Cyster, J. G. EBI2 mediates B cell segregation between the outer and centre follicle. *Nature* **460**, 1122–1126 (2009).
- Gatto, D., Paus, D., Basten, A., Mackay, C. R. & Brink, R. Guidance of B cells by the orphan G protein-coupled receptor EBI2 shapes humoral immune responses. *Immunity* **31**, 259–269 (2009).
- Heinig, M. *et al.* A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature* **467**, 460–464 (2010).
- Russell, D. W. The enzymes, regulation, and genetics of bile acid synthesis. *Annu. Rev. Biochem.* **72**, 137–174 (2003).
- Lathe, R. Steroid and sterol 7-hydroxylation: ancient pathways. *Steroids* **67**, 967–977 (2002).
- Russell, D. W. Oxysterol biosynthetic enzymes. *Biochim. Biophys. Acta* **1529**, 126–135 (2000).
- Bensinger, S. J. *et al.* LXR signaling couples sterol metabolism to proliferation in the acquired immune response. *Cell* **134**, 97–111 (2008).
- Gauwerky, C. E., Hoxie, J., Nowell, P. C. & Croce, C. M. Pre-B-cell leukemia with a t(8;14) and a t(14;18) translocation is preceded by follicular lymphoma. *Oncogene* **2**, 431–435 (1988).
- Reed, J. C. & Tanaka, S. Somatic point mutations in the translocated *bcl-2* genes of non-Hodgkin's lymphomas and lymphocytic leukemias: implications for mechanisms of tumor progression. *Leuk. Lymphoma* **10**, 157–163 (1993).
- Novershtern, N. *et al.* Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* **144**, 296–309 (2011).
- Rubic, T. *et al.* Triggering the succinate receptor GPR91 on dendritic cells enhances immunity. *Nature Immunol.* **9**, 1261–1269 (2008).
- Bauman, D. R. *et al.* 25-Hydroxycholesterol secreted by macrophages in response to Toll-like receptor activation suppresses immunoglobulin A production. *Proc. Natl Acad. Sci. USA* **106**, 16764–16769 (2009).
- Diczfalusy, U. *et al.* Marked upregulation of cholesterol 25-hydroxylase expression by lipopolysaccharide. *J. Lipid Res.* **50**, 2258–2264 (2009).
- Honda, A. *et al.* Cholesterol 25-hydroxylation activity of CYP3A. *J. Lipid Res.* doi:10.1194/jlr.M014084 (2011).
- Li, X. *et al.* Biosynthesis of the regulatory oxysterol, 5-cholesten-3 $\beta$ ,25-diol 3-sulfate, in hepatocytes. *J. Lipid Res.* **48**, 2587–2596 (2007).
- Lund, E., Bjorkhem, I., Furster, C. & Wikvall, K. 24-, 25- and 27-hydroxylation of cholesterol by a purified preparation of 27-hydroxylase from pig liver. *Biochim. Biophys. Acta* **1166**, 177–182 (1993).
- Dulos, J., van der Vleuten, M. A., Kavelaars, A., Heijnen, C. J. & Boots, A. M. CYP7B expression and activity in fibroblast-like synoviocytes from patients with rheumatoid arthritis: regulation by proinflammatory cytokines. *Arthritis Rheum.* **52**, 770–778 (2005).
- Wang, J. H. & Tuohimaa, P. Regulation of cholesterol 25-hydroxylase expression by vitamin D3 metabolites in human prostate stromal cells. *Biochem. Biophys. Res. Commun.* **345**, 720–725 (2006).
- Park, K. & Scott, A. L. Cholesterol 25-hydroxylase production by dendritic cells and macrophages is regulated by type I interferons. *J. Leukoc. Biol.* **88**, 1081–1087 (2010).
- Liu, C. *et al.* Oxysterols direct B-cell migration through EBI2. *Nature* doi:10.1038/10226 (2011).
- Li, D. & Spencer, T. A. Synthesis of 7 $\alpha$ -hydroxy derivatives of regulatory oxysterols. *Steroids* **65**, 529–535 (2000).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We would like to thank the following scientists for their support: F. Su and J.-L. Vincent (Bruxelles Erasme Hospital) for access to tissue samples from the peritonitis sheep sepsis model; D. W. Russell (University of Texas Southwestern) for access to the *Ch25h*<sup>-/-</sup> mouse strain; J. Carballido and T. Rubic (all Novartis) for providing an EBV-infected B-cell line and help in preparation of mouse bone-marrow-derived dendritic cells; J. C. Reed (Burnham Institute for Medical Research) for the RS11846 cells; A.-G. Letombe (Euroscreen), A. Feige and N. Beluch for technical assistance; and J. Hamon, C. Bauer, A. Glaenzel, P. Bross, V. Techer-Etienne (all Novartis) and RC Tritic Ltd (Teufen, Switzerland) for support in the development of a radioligand binding assay.

**Author Contributions** S.H., S.N., M.-O.R., M.D. and K.S. initiated the project and developed strategy. S.H., S.N. and M.-O.R. purified activity from sheep tissue extracts. W.S., D.-I.K., C.Y.C. and P.G.S. carried out corresponding work on the pig tissue. J.Z., R.K., I.C., R.A.K., R.F. and E.C.P. conducted mass-spectrometry and NMR experiments. S.L., M.V., I.P., D.G., D.N., W.L., A.K., Y.A.C., B.U.B., K.S. and A.W.S. designed and conducted pharmacological characterization of the receptor and bioactivity assays. S.R., C.M., F.G., Ca.S., W.S., C.Y.C., S.H. and A.W.S. proposed and/or synthesized various oxysterol ligands. T.S., Ch.S. and B.W. supported *in vivo* experiments with mouse models. T.Y., J.P.P., L.M.K. and J.G.C. developed the bioassay, conducted the experiments in CH25H-deficient and transduced mice and performed the B-cell chemotaxis and immunization studies. A.W.S. wrote the first draft of the manuscript, which was completed by S.H., J.Z., D.N., B.W., S.R., J.G.C., C.Y.C., K.S. and commented on by all authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to A.W.S. ([andreas.sailer@novartis.com](mailto:andreas.sailer@novartis.com)).

# The crystal structure of GXGD membrane protease FlaK

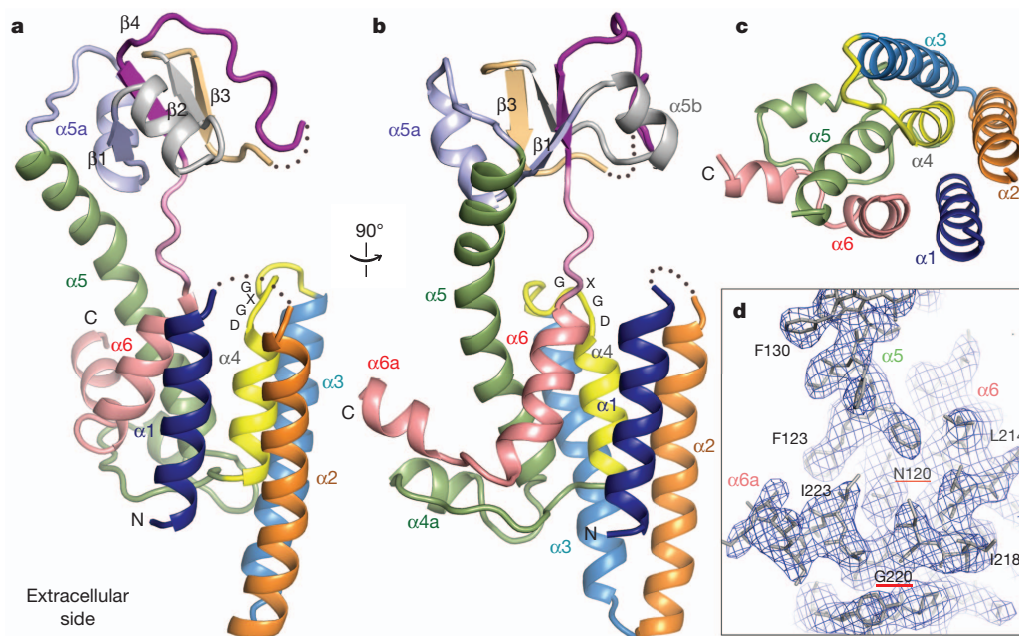
Jian Hu<sup>1</sup>, Yi Xue<sup>1</sup>, Sangwon Lee<sup>1</sup> & Ya Ha<sup>1</sup>

The GXGD proteases are polytopic membrane proteins with catalytic activities against membrane-spanning substrates that require a pair of aspartyl residues<sup>1–4</sup>. Representative members of the family include preflagellin peptidase, type 4 prepilin peptidase, presenilin and signal peptide peptidase. Many GXGD proteases are important in medicine. For example, type 4 prepilin peptidase may contribute to bacterial pathogenesis<sup>5–7</sup>, and mutations in presenilin are associated with Alzheimer's disease<sup>8–10</sup>. As yet, there is no atomic-resolution structure in this protease family. Here we report the crystal structure of FlaK, a preflagellin peptidase from *Methanococcus maripaludis*, solved at 3.6 Å resolution. The structure contains six transmembrane helices. The GXGD motif and a short transmembrane helix, helix 4, are positioned at the centre, surrounded by other transmembrane helices. The crystal structure indicates that the protease must undergo conformational changes to bring the GXGD motif and a second essential aspartyl residue from transmembrane helix 1 into close proximity for catalysis. A comparison of the crystal structure with models of presenilin derived from biochemical analysis reveals three common transmembrane segments that are similarly arranged around the active site. This observation reinforces the idea that the prokaryotic and human proteases are evolutionarily related<sup>11,12</sup>.

The crystal structure presented here provides a framework for understanding the mechanism of the GXGD proteases, and may facilitate the rational design of inhibitors that target specific members of the family.

Archaeal preflagellins and bacterial type-4 prepilins, both of which are type-II (N<sub>in</sub>-C<sub>out</sub>) membrane proteins<sup>13,14</sup>, are synthesized with short, positively charged leader peptides<sup>5,15</sup>. They are cleaved in the membrane, at a site a few residues upstream of the hydrophobic membrane-spanning sequence, by preflagellin peptidase (PFP) and type-4 prepilin peptidase (TFPP), respectively, before being secreted and incorporated into the mature flagellum or type-4 pilus (Supplementary Fig. 1a). We have crystallized FlaK, a prototypic archaeal PFP from *M. maripaludis*<sup>16</sup>. The membrane protease maintained a native-like conformation throughout crystallization because both the crystallization drop and the dissolved crystals showed robust enzymatic activity (Supplementary Fig. 1b). The structure was determined by single-wavelength anomalous dispersion using a Se-Met-substituted crystal (Supplementary Fig. 2 and Supplementary Table 1).

The crystal structure shows that FlaK contains two compactly folded domains: a mostly  $\alpha$ -helical membrane-spanning domain and a soluble domain with four anti-parallel  $\beta$ -strands (Fig. 1; the soluble domain is disordered in one of the two FlaK molecules in the asymmetric unit).

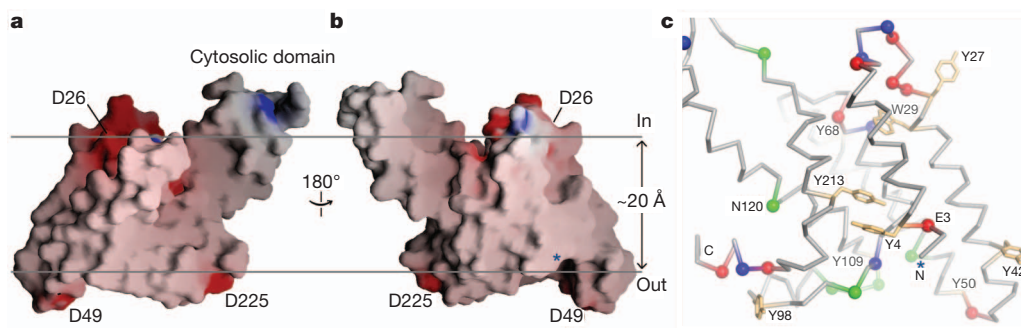


**Figure 1 | The structure of FlaK.** Shown here are cartoon representations of molecule A, one of the two FlaK molecules in the asymmetric unit. In molecule B, the soluble domain cannot be completely traced owing to disorder. These illustrations, as well as those in Figs 2c, 3a, b and 4b, c, were generated by PyMOL (<http://www.pymol.org>). **a, b**, Two views of the molecule from the side.

The secondary structural elements and the GXGD motif are labelled. **c**, A view from the cytosolic side of the membrane. For clarity, the soluble domain (including part of  $\alpha 5$ ) is removed. **d**, A portion of the final  $2F_o - F_c$  electron density map (contoured at  $1\sigma$  level) showing the position of  $\alpha 6a$  relative to TM helix  $\alpha 5$ .

<sup>1</sup>Department of Pharmacology, Yale School of Medicine, 333 Cedar Street, New Haven, Connecticut 06520, USA.





**Figure 2 | FlaK is tilted in the membrane.** Molecule B is shown in this figure because its TM region is better defined in the electron density map, and does not contain any breaks. **a, b**, The molecular surface, colour-coded by electrostatic potential. The two horizontal lines roughly mark the hydrophobic belt around the membrane protease. The N terminus of the protein is labelled

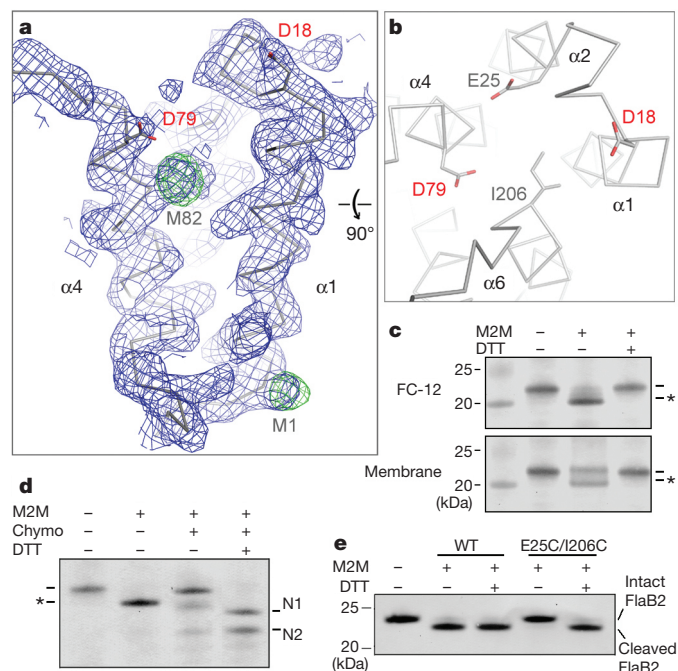
Previous predictions correctly located all six transmembrane (TM) segments<sup>4</sup>, but the crystal structure is more complex (Supplementary Fig. 3a). One major deviation from the prediction occurs around TM4 (yellow in Fig. 1). The hydrophilic loop between TM3 and TM4 does not protrude into the cytoplasm as predicted; instead, it lowers towards the centre of the bundle of TM helices, and is followed by a short TM helix,  $\alpha 4$ . The last TM helix ( $\alpha 6$ ) is also short (pink in Fig. 1), and seems unable to cross the lipid bilayer completely. The protein segments immediately after  $\alpha 4$  and  $\alpha 6$  form an unusual structure that protrudes sideways from the base of the TM helices. This feature does not seem to be an artefact of crystallization. A comparison between the two copies of FlaK in the asymmetric unit shows that the unusual structure is identically positioned, despite the fact that it is involved in different crystal packing interactions (Supplementary Fig. 3b). The amphipathic nature of the structure is also consistent with its position next to the TM helices, and with the possibility that it may interact peripherally with the membrane. In the big loop between  $\alpha 4$  and  $\alpha 5$  (including  $\alpha 4a$ ), and in the carboxy-terminal segment (including  $\alpha 6a$ ), all the polar side chains point downwards away from the membrane, whereas most hydrophobic side chains point up or sideways to interact either with the TM helices, or with lipids that surround the helices. Furthermore, there is a conserved asparagine on  $\alpha 5$  (Asn 120), the side chain of which points outwards to form a hydrogen bond with the carbonyl oxygen of Gly 220 from the extended segment between  $\alpha 6$  and  $\alpha 6a$  (Fig. 1d). If the carboxy-terminus of the protein were positioned elsewhere, Asn 120 would become unfavourably exposed to the lipid.

To accommodate the unusual peripheral structure ( $\alpha 4a$  and  $\alpha 6a$ ), and the short TM helix  $\alpha 6$ , the other TM helices must be tilted in the membrane. The tilting is required to avoid positioning charged groups such as Asp 26, Asp 49 and Asp 225 in the hydrophobic region of the membrane (Fig. 2). The tilting also makes  $\alpha 6$  roughly perpendicular to the membrane plane, enabling it to go through the lipid bilayer with the shortest distance. A more thorough examination of the distribution of amino acids in the TM region supports the tilted model. As shown in Fig. 2c, the lower boundary of the membrane is roughly marked by a thin belt of acidic residues (red), basic residues (blue), asparagines and glutamines (green). Four tyrosine residues (Tyr 42, Tyr 50, Tyr 98 and Tyr 109) are also found within this belt. Tyrosine and tryptophan often cluster to the interface between water and lipid<sup>17</sup>. The upper boundary of the membrane probably corresponds to a plane that goes through Tyr 27, Trp 29 and Tyr 68. Although Glu 3, Tyr 4, Asn 120 and Tyr 213 appear between the two boundaries, a closer inspection shows that the polar groups on their side chains are not directly exposed to the lipid. According to the tilted model, the membrane has to be constricted around the protease. This feature was previously thought to be unique to the intramembrane proteases<sup>18–20</sup>.

The two aspartyl residues that are essential for catalysis (Asp 18 and Asp 79) are located at the ends of TM helices  $\alpha 1$  and  $\alpha 4$ , respectively.

by a blue asterisk. **c**, The  $C_{\alpha}$  trace of the TM region with the same orientation as in panel **b**. Red spheres, negatively charged residues; blue spheres, positively charged residues; green spheres, asparagines and glutamines. Tyrosines and tryptophans are shown as orange stick models (undefined side chains are not shown).

Despite the relatively low resolution at which the crystal structure was solved,  $\alpha 1$  and  $\alpha 4$  are clearly defined in the electron density map, and the experimentally determined selenium sites in both helices are consistent with the register of protein sequence with the density (Fig. 3a). The positions of  $\alpha 1$  and  $\alpha 4$  are almost identical in the two independently modelled FlaK molecules in the asymmetric unit of the crystal (Supplementary Fig. 3b). The spatial relationship between the two

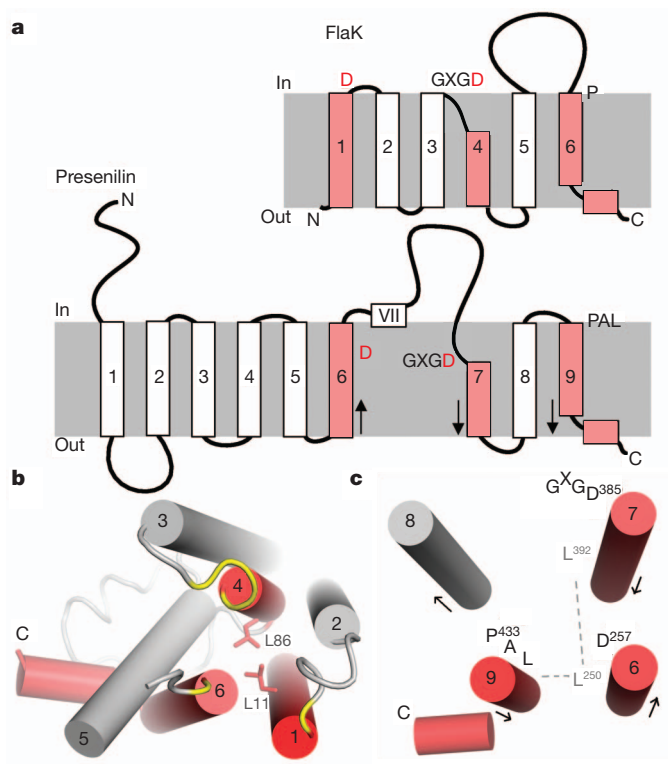


**Figure 3 | The uncoupling between Asp 18 and Asp 79.** **a**, The  $2F_o - F_c$  map (contoured at  $1\sigma$  level; blue) and the anomalous Fourier map (contoured at  $4\sigma$  levels; green) around  $\alpha 1$  and  $\alpha 4$ . **b**, A view from the cytosolic side of the membrane. The distance from  $C_{\alpha 1}$  of Ile 206 to  $C_{\alpha 7}$  of Glu 25 is  $\sim 6$  Å. **c**, Crosslinking causes the E25C/I206C mutant protein to migrate faster in the SDS–polyacrylamide gel (to a position marked by the asterisk); adding DTT breaks the disulphide linkage. FC-12, foscholine-12. The left lane is a molecular weight marker. **d**, Partial proteolysis confirms the covalent linkage between Cys 25 and Cys 206. Chymotrypsin (chymo) cleaves FlaK twice between TM domains 5 and 6, generating two N-terminal fragments (N1 and N2). For crosslinked mutant protein, the N-terminal fragments (containing Cys 25) remain attached to the C-terminal fragment (containing Cys 206). **e**, Crosslinking between Cys 25 and Cys 206 renders the protease completely inactive; treatment with DTT fully restores activity. Wild-type (WT) FlaK is not affected by M2M and DTT. The antibody used in the western blots (panels **c–e**) detects the N-terminal His tag of FlaK and the C-terminal His tag of FlaB2 (substrate).

helices is, however, surprising from a functional point of view because it creates a 12 Å gap between Asp 18 and Asp 79. To investigate whether the FlaK crystal structure represents a non-active conformation of the protease, we introduced a pair of cysteines (E25C and I206C) to the two TM helices ( $\alpha 2$  and  $\alpha 6$ ) that are on the opposite side of the gap (Fig. 3b). The double mutant (E25C/I206C) was proteolytically active, indicating that its conformation was not markedly perturbed by the mutations. As shown in Fig. 3c, d, the two cysteines can be readily crosslinked by 1,2-ethanedithiol dimethanesulphonate (M2M). Given its short length, the crosslinker has to lie inside the gap between  $\alpha 1$  and  $\alpha 4$  to bridge Cys 25 and Cys 206. Taken together, these results indicate that, in the absence of substrate, the membrane protease can adopt an inactive conformation in which the two catalytic aspartyl residues are structurally uncoupled. Consistent with this idea, we found that crosslinking between  $\alpha 2$  and  $\alpha 6$ , which prevents movement of Asp 18 and Asp 79 towards each other, completely eliminated protease activity (Fig. 3e). Breaking the crosslinking disulphide bonds with dithiothreitol (DTT) fully restored protease activity. The observation that crosslinking in the membrane fraction is less complete than in detergent solutions indicates that FlaK may assume additional conformations in the lipid bilayer (Fig. 3c). The behaviour of FlaK is similar to that of presenilin in these regards: the human protease also switches between at least two conformations and in one conformation the two catalytic aspartates do not closely oppose each other<sup>21,22</sup>.

Three regions in the FlaK sequence are highly conserved (Supplementary Fig. 4). The first two are centred on Asp 18 and on the GXGD motif (Asp 79). The third region, which corresponds to the sequence around the amino terminus of TM helix  $\alpha 6$ , is shown by the crystal structure to be near the active site as well. The roles of individual residues from the conserved regions in the protease mechanism were probed by mutagenesis (Supplementary Fig. 5). Three main lessons can be learned from this analysis. First, Glu 23, Glu 25 and Asp 26, the only three acidic residues around the active site, do not seem to be essential for the binding of the positively charged leader peptide. Second, among the three glycine residues in the GXGD motif, Gly 76 is the most critical for function. Because Gly 76 is not closely packed against other residues, the large reduction of activity in the G76A mutant probably results from a partial loss of backbone flexibility in this important region. Third, most mutations from Pro 201 to Pro 208 produce a small effect; P208A, however, is different in that it markedly enhances enzymatic activity. Pro 208 is packed against TM helix  $\alpha 5$ , and points away from the active site. If the conformational change in the protease involves movement of TM helices  $\alpha 1$ ,  $\alpha 4$  and  $\alpha 6$  (Fig. 3b), then the altered packing between  $\alpha 6$  and  $\alpha 5$  might have facilitated such a change.

Presenilin has nine TM segments<sup>23,24</sup> (Fig. 4a). The last four segments (TM6–TM9) of presenilin share homology with signal peptide peptidase<sup>25</sup>, and are thought to constitute the core of the protease<sup>25</sup>. Previous cysteine-scanning mutagenesis and crosslinking experiments indicated that the active site of presenilin is housed in an open hydrophilic cavity<sup>21,26</sup>, surrounded by the two TM segments (TM6 and TM7) that carry the catalytic aspartates (Asp 257 and Asp 385), and by the C-terminal TM9, which bears a conserved Pro-Ala-Leu motif<sup>22,27</sup>. The crystal structure of FlaK now shows that the active site of the prokaryotic protease has a similar architecture. As illustrated in Fig. 4a, the TM segments TM1, TM4 and TM6 of FlaK would be equivalent to presenilin's TM6, TM7 and TM9. In FlaK, there is a highly conserved leucine (Leu 11) seven residues upstream of the catalytic Asp 18 on TM1 (Supplementary Fig. 4). Leu 11 is packed against Leu 86 from TM4 (Fig. 4b). Leu 86, which is also conserved, is seven residues downstream of the catalytic Asp 79. The residues in presenilin corresponding to Leu 11 and Leu 86 would be Leu 250 and Leu 392, respectively (seven residues away from the catalytic Asp 257 and Asp 385). Both Leu 250 and Leu 392 are highly conserved<sup>28</sup>. The fact that Asp 257 and Asp 385 are close to each other indicates that Leu 250 and Leu 392, which are on the same side of the helices as the aspartates, may also



**Figure 4 | Structural comparison between FlaK and presenilin-1.**

**a**, Topology diagrams of FlaK and presenilin. The three key TM segments are highlighted in red. The grey boxes represent membrane. Arrows indicate the direction of the helices (N to C termini). **b**, FlaK viewed from the cytosolic side of the membrane. Asp 18, the GXGD motif and the conserved Pro 204 are shown in yellow. The two conserved leucines are shown as stick models. **c**, Packing of the three key TM helices in presenilin. This model is similar to the one proposed in ref. 27, but is the mirror image of another model<sup>22</sup>. Leu 250 from TM6 is hypothesized to mediate packing with TM7 (through Leu 392) and TM9.

interact (Fig. 4c). In FlaK, Leu 11 makes additional contact with Tyr 213 from TM6. In presenilin, Leu 250 also seems to interact with the last TM segment because it can be readily crosslinked to many positions on TM9<sup>27</sup>. Besides the packing of the three key helices, presenilin and FlaK share other features. TM7 of presenilin, like its counterpart in FlaK, also seems to contain two structural elements: an exposed region bearing the GXGD motif and a short, tightly packed hydrophobic helix<sup>26</sup>. The last TM segments in both proteases have conserved proline residues near the N terminus, and are followed by an amphipathic helix that interacts peripherally with the membrane<sup>27</sup>. It is important to note that presenilin lacks the equivalent of FlaK's TM segments 2 and 3. Therefore, other TM helices may join the central three TM segments to complete the active site. The hydrophobic domain VII, which undergoes endo-proteolysis<sup>1</sup>, must also be bound initially at the active site<sup>22</sup>.

The structure of FlaK's active site is fundamentally different from that of pepsin, a classic aspartyl protease<sup>29</sup>, in that it lacks an internal two-fold symmetry and its two catalytic aspartyl residues are not rigidly fixed. Membrane proteases have evolved unique mechanisms to conduct catalysis inside lipid bilayers. For example, rhomboid serine proteases use a surface cap to control access to a preformed and membrane-embedded Ser-His catalytic dyad<sup>30</sup>. The uncoupling between two catalytic aspartyl residues, indicated by the crystal structure described here and by earlier biochemical studies<sup>21,22</sup>, may represent a general mechanism that is widely adopted by the GXGD proteases. Such an uncoupling mechanism could potentially have an important role in regulating catalysis.

## METHODS SUMMARY

MmarC6\_0338 (encoding FlaK) was amplified by PCR from the genomic DNA of *M. maripaludis* strain C6, cloned into pET-28a and expressed in *Escherichia coli* Rosetta 2 (DE3) cells (Novagen). The Se-Met-substituted membrane protein was extracted in foscholine-12 (Anatrace) and purified using metal-affinity and size-exclusion columns. The concentrated protein ( $\sim 10 \text{ mg ml}^{-1}$ ) was extensively dialysed against a buffer containing 20 mM HEPES (pH 7.3), 100 mM NaCl, 0.06% Cymal-6 (Anatrace) and 1 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP). Single crystals were prepared by the sitting-drop method, in which 1  $\mu\text{l}$  of protein solution was mixed with 1  $\mu\text{l}$  of well solution containing 30% PEG 300, 50 mM glycine (pH 9.5) and 100 mM NaCl. Many crystals were screened at the national synchrotron light source (NSLS, X25 and X29) and at the advanced photon source (APS, 24-ID-C and E) before a final data set at 3.6 Å resolution was collected from a single Se-Met-substituted crystal at the selenium peak wavelength, which was used for both phase determination and refinement. The details of protein purification, crystal structure determination, protease activity assay and crosslinking experiments are described in the Methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 21 March; accepted 18 May 2011.**

**Published online 17 July 2011.**

- Wolfe, M. S. *et al.* Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and  $\gamma$ -secretase activity. *Nature* **398**, 513–517 (1999).
- LaPointe, C. F. & Taylor, R. K. The type 4 prepilin peptidases comprise a novel family of aspartic acid proteases. *J. Biol. Chem.* **275**, 1502–1510 (2000).
- Weihsen, A., Binns, K., Lemberg, M. K., Ashman, K. & Martoglio, B. Identification of signal peptide peptidase, a presenilin-type aspartic protease. *Science* **296**, 2215–2218 (2002).
- Bardy, S. L. & Jarrell, K. F. Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon *Methanococcus voltae*. *Mol. Microbiol.* **50**, 1339–1347 (2003).
- Lory, S. & Strom, M. S. Structure-function relationship of type-IV prepilin peptidase of *Pseudomonas aeruginosa*—a review. *Gene* **192**, 117–121 (1997).
- Craig, L., Pique, M. E. & Tainer, J. A. Type IV pilus structure and bacterial pathogenicity. *Nature Rev. Microbiol.* **2**, 363–378 (2004).
- Sandkvist, M. Type II secretion and pathogenesis. *Infect. Immun.* **69**, 3523–3535 (2001).
- Selkoe, D. J. & Wolfe, M. S. Presenilin: running with scissors in the membrane. *Cell* **131**, 215–221 (2007).
- Jorissen, E. & De Strooper, B.  $\gamma$ -secretase and the intramembrane proteolysis of Notch. *Curr. Top. Dev. Biol.* **92**, 201–230 (2010).
- Brouwers, N., Sleegers, K. & Van Broeckhoven, C. Molecular genetics of Alzheimer's disease: an update. *Ann. Med.* **40**, 562–583 (2008).
- Steiner, H. *et al.* Glycine 384 is required for presenilin-1 function and is conserved in bacterial polytopic aspartyl proteases. *Nature Cell Biol.* **2**, 848–851 (2000).
- Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. & Barrett, A. J. MEROPS: the peptidase database. *Nucleic Acids Res.* **36**, D320–D325 (2008).
- Francetic, O., Buddelmeijer, N., Lewenza, S., Kumamoto, C. A. & Pugsley, A. P. Signal recognition particle-dependent inner membrane targeting of the PulG Pseudopilin component of a type II secretion system. *J. Bacteriol.* **189**, 1783–1793 (2007).
- Bayley, D. P. & Jarrell, K. F. Overexpression of *Methanococcus voltae* flagellin subunits in *Escherichia coli* and *Pseudomonas aeruginosa*: a source of archaeal preflagellin. *J. Bacteriol.* **181**, 4146–4153 (1999).
- Kalmokoff, M. L., Karnauchow, T. M. & Jarrell, K. F. Conserved N-terminal sequences in the flagellins of archaeobacteria. *Biochem. Biophys. Res. Commun.* **167**, 154–160 (1990).
- Bardy, S. L. & Jarrell, K. F. FlaK of the archaeon *Methanococcus maripaludis* possesses preflagellin peptidase activity. *FEMS Microbiol. Lett.* **208**, 53–59 (2002).
- Killian, J. A. & von Heijne, G. How proteins adapt to a membrane–water interface. *Trends Biochem. Sci.* **25**, 429–434 (2000).
- Wang, Y., Maegawa, S., Akiyama, Y. & Ha, Y. The role of L1 loop in the mechanism of rhomboid intramembrane protease GlpG. *J. Mol. Biol.* **374**, 1104–1113 (2007).
- Bondar, A. N., del Val, C. & White, S. H. Rhomboid protease dynamics and lipid interactions. *Structure* **17**, 395–405 (2009).
- Ha, Y. Structure and mechanism of intramembrane protease. *Semin. Cell Dev. Biol.* **20**, 240–250 (2009).
- Tolia, A., Chavez-Gutierrez, L. & De Strooper, B. Contribution of presenilin transmembrane domains 6 and 7 to a water-containing cavity in the  $\gamma$ -secretase complex. *J. Biol. Chem.* **281**, 27633–27642 (2006).
- Tolia, A., Horre, K. & De Strooper, B. Transmembrane domain 9 of presenilin determines the dynamic conformation of the catalytic site of  $\gamma$ -secretase. *J. Biol. Chem.* **283**, 19793–19803 (2008).
- Laudon, H. *et al.* A nine-transmembrane domain topology for presenilin 1. *J. Biol. Chem.* **280**, 35352–35360 (2005).
- Spasic, D. *et al.* Presenilin-1 maintains a nine-transmembrane topology throughout the secretory pathway. *J. Biol. Chem.* **281**, 26569–26577 (2006).
- Narayanan, S., Sato, T. & Wolfe, M. S. A C-terminal region of signal peptide peptidase defines a functional domain for intramembrane aspartic protease catalysis. *J. Biol. Chem.* **282**, 20172–20179 (2007).
- Sato, C., Morohashi, Y., Tomita, T. & Iwatsubo, T. Structure of the catalytic pore of  $\gamma$ -secretase probed by the accessibility of substituted cysteines. *J. Neurosci.* **26**, 12081–12088 (2006).
- Sato, C., Takagi, S., Tomita, T. & Iwatsubo, T. The C-terminal PAL motif and transmembrane domain 9 of presenilin 1 are involved in the formation of the catalytic pore of the  $\gamma$ -secretase. *J. Neurosci.* **28**, 6264–6271 (2008).
- Ponting, C. P. *et al.* Identification of a novel family of presenilin homologues. *Hum. Mol. Genet.* **11**, 1037–1044 (2002).
- Davies, D. R. The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 189–215 (1990).
- Wang, Y. & Ha, Y. Open-cap conformation of intramembrane protease GlpG. *Proc. Natl Acad. Sci. USA* **104**, 2098–2102 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank A. Héroux, H. Robinson and A. Soares at NSLS, and J. Schuermann at APS NE-CAT for their help during data collection. X-ray diffraction data were measured at beamlines X25 and X29 at NSLS, and at 24-ID-C and 24-ID-E at APS. Financial support was principally from the US Department of Energy and from the National Institutes of Health. This work was supported by a New Scholar Award in Aging from the Ellison Medical Foundation (to Y.H.), a gift from the Neuroscience Education and Research Foundation (to Y.H.) and a pilot grant from Yale's programme in Cellular Neuroscience, Neurodegeneration, and Repair (CNR) (to Y.H.).

**Author Contributions** J.H. and Y.X. purified and characterized FlaK in various detergents. J.H. obtained the high-resolution crystals of FlaK. J.H., Y.X. and Y.H. solved the crystal structure. Y.H., Y.X. and J.H. wrote the paper. Y.X. and S.L. screened many constructs and performed the initial biochemical and functional characterizations.

**Author Information** The atomic coordinates of FlaK and structure factors have been deposited in the Protein Data Bank under accession code 3SOX. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to Y.H. ([ya.ha@yale.edu](mailto:ya.ha@yale.edu)).



## METHODS

**Protein expression and purification.** Many PFPs and TFPPs were screened for crystallization. Most of the genes were amplified by PCR from genomic DNAs purchased from ATCC. The genes for *Pseudomonas aeruginosa* PilD/PilA, *E. coli* BfpP and *Dichelobacter nodosus* FimP were gifts from S. Lory, M. Donnenberg and J. Rood, respectively. MmarC6\_0338 (encoding FlaK) was amplified from the genomic DNA of *M. maripaludis* strain C6, and cloned into pET-28a. To facilitate removal of the N-terminal His tag, a Gly-Ser-Gly-Ser sequence was inserted between the thrombin site and FlaK sequence. Mvol\_1295 (encoding FlaB2) was amplified from the genomic DNA of *Methanococcus voltae* strain A3 and cloned into pET-43b with a C-terminal His tag<sup>16</sup>. FlaK was expressed in *E. coli* Rosetta 2 (DE3) cells (Novagen), grown in Luria broth. FlaB2 was similarly over-expressed in C43(DE3) cells (Lucigen). To generate Se-Met FlaK, cells were cultured at 37 °C in M9 minimum media supplemented with Se-Met, then induced by 1 mM  $\beta$ -D-thiogalactopyranoside (IPTG) at an absorbance at 600 nm ( $A_{600}$ ) of 0.6, and grown at 20 °C for 16 h before collection. Cytoplasmic membranes were prepared by the spheroplast method<sup>31</sup> and suspended in a buffer containing 50 mM sodium phosphate (pH 7.2), 500 mM NaCl, 5 mM  $\beta$ -mercaptoethanol and a cocktail of complete protease inhibitors (tablet without EDTA, Roche Diagnostics). For solubilization, powder of foscholine-12 (Anatrace) was added to the membrane suspension to achieve a final concentration of 1% (w/v). The His-tagged protein was eluted from a TALON metal affinity column (Clontech) in 50 mM sodium phosphate (pH 7.2), 500 mM NaCl, 200 mM imidazole, 5 mM  $\beta$ -mercaptoethanol and 0.1% foscholine-12. After passing through a Sephadex G-25 desalting column, the sample was cleaved by thrombin overnight at 22 °C. Finally, the protein was loaded onto a Superdex-200 column (GE Healthcare) equilibrated with 20 mM HEPES (pH 7.3), 100 mM NaCl, 1 mM TCEP and 0.1% foscholine-12. The peak fraction was pooled, concentrated to 10 mg ml<sup>-1</sup> and dialysed against 20 mM HEPES (pH 7.3), 100 mM NaCl, 1 mM TCEP and 0.06% Cymal-6 (Anatrace) at 4 °C for 8 days. About 3 mg of purified membrane protein could be obtained for crystallization trials from 1 litre of bacterial culture.

**Crystallization and structure determination.** The sitting-drop method was used to prepare Se-Met FlaK crystals: 1  $\mu$ l of protein solution (4 mg ml<sup>-1</sup>; the lower concentration is due to precipitation during dialysis) was mixed with 1  $\mu$ l of well solution containing 30% PEG 300, 50 mM glycine (pH 9.5) and 100 mM NaCl. Needle-shaped crystals usually appeared in 2 days at 22 °C and grew to full size in 1 week. The crystals were dehydrated by equilibrating for 24 h against a well solution containing 40% PEG 300, before direct flash-freezing in liquid nitrogen. Screening and data collection were performed at the national synchrotron light source (X25 and X29) and at the advanced photon source (24-ID-C and E). All diffraction data were processed by HKL2000 (ref. 32). The structure was determined by single-wavelength anomalous dispersion<sup>33</sup> using a highly redundant data set which was generated by merging four data sets collected at four different spots on a single Se-Met crystal at 24-ID-C. The same data set was used in refinement (Supplementary Table 1). The selenium sites and the initial phases were determined by hkl2map<sup>34</sup>. The experimental electron density map confirmed the presence of two FlaK molecules in the asymmetric unit, and clearly showed all the TM helices (Supplementary Fig. 2). The soluble domain in molecule A was visible but could not yet be traced; the soluble domain in molecule B was mostly missing. Averaging the TM regions of the two molecules by dm<sup>35</sup> improved the clarity of the map. Modelling of the polypeptide chains using O<sup>36</sup> was assisted by known Se sites (Supplementary Fig. 6). After rounds of model building and refinement by CNS<sup>37</sup>, the phases were sufficiently improved to allow complete tracing of the soluble domain in molecule A. The final model was refined by CNS and

refmac5 (ref. 38). The electrostatic potential surfaces shown in Fig. 2a, b were generated by GRASP<sup>39</sup>.

**FlaK activity assay.** The enzymatic activity of FlaK was measured according to ref. 4. In brief, membranes containing overexpressed FlaK or FlaB2 were prepared using the spheroplast method, and were re-suspended in phosphate buffer. The membrane fractions were then mixed and the reaction, at 22 °C, was initiated by adding a  $\times 5$  reaction buffer containing 2.5% Triton X-100 and 100 mM HEPES (pH 7.3). The reaction was stopped by mixing with SDS–polyacrylamide gel electrophoresis sample-loading buffer. The reaction mixture was examined by western blot using an anti-His-tag antibody (Calbiochem). The purified FlaK in detergent solutions was assayed similarly. The two assays shown in Supplementary Fig. 1b were conducted at 22 °C for 120 min and 90 min, respectively: in the top panel, a large amount of protease (as indicated by the protease band) and a long incubation time were used to exclude the possibility of residual enzymatic activity in the asparagine mutants. In Supplementary Fig. 5, a shorter assay time (45 min) and a smaller amount of protease were used so that both the intact and processed FlaB2 are visible: in this setting, the amount of intact substrate is a good indicator of the reaction rate. The amount of protease used in the assay was reflected in the lower control panel, where the loading (of FlaK alone) was ten times that used in the assay, to increase the visibility of the protease. The same amount of protease was used in the assay shown in Fig. 3e, but the reaction time was longer (120 min).

**Chemical crosslinking.** The membrane preparation containing the E25C/I206C double mutant was suspended in a buffer containing 50 mM sodium phosphate (pH 8.0) and 100 mM NaCl. Crosslinking was performed either directly in the membrane suspension or in a foscholine-12-solubilized membrane fraction (0.2% foscholine-12), by treating the protein with 2 mM M2M (Toronto Research Chemicals Inc.) at 22 °C for 1 h. M2M has a spacer arm length of 5.2 Å<sup>40,41</sup>. The reaction was stopped by adding an equal volume of 40 mM N-ethylmaleimide in 200 mM HEPES (pH 7.3). DTT (final concentration 50 mM) was used to break the crosslinking disulphide bond.

- Kaback, H. R. & Stadtman, E. R. Proline uptake by an isolated cytoplasmic membrane preparation of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **55**, 920–927 (1966).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Wang, B. C. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115**, 90–112 (1985).
- Pape, T. & Schneider, T. R. HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Cryst.* **37**, 843–844 (2004).
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Brünger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
- Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **374**, 300–321 (2003).
- Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).
- Loo, T. W. & Clarke, D. M. Determining the dimensions of the drug-binding domain of human P-glycoprotein using thiol cross-linking compounds as molecular rulers. *J. Biol. Chem.* **276**, 36877–36880 (2001).
- Green, N. S., Reisler, E. & Houk, K. N. Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. *Protein Sci.* **10**, 1293–1304 (2001).

# CAREERS

**ADVICE FORUM** Discuss all aspects of scientific careers [go.nature.com/lm1x4t](http://go.nature.com/lm1x4t)

**BLOG** The latest news and information about research jobs [go.nature.com/z8g4a7](http://go.nature.com/z8g4a7)

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)

IMAGES.COM/CORE/iS



## GRADUATE STUDENTS

# Aspirations and anxieties

*Nature's international student survey reveals changing career preferences — and a need for inspiring mentors.*

BY GENE RUSSO

Life as a graduate student can mean hours of daily toil, little social contact and no guarantee that all that work will lead to a job. But it can also offer intellectual stimulation, independent projects that nurture a love of discovery and the development of a skill set that opens a host of science-related opportunities for a budding scientist.

No wonder, then, that a survey of graduate

students suggests many are ambivalent about their work and their future. The results imply that doubts grow as students advance towards completing their degrees, and suggest that students experience deficiencies in career planning and advice from their supervisors (see 'Changing attitudes').

About 5,000 graduate students from dozens of countries responded to the survey, which *Nature* publicized through its e-mail lists and website, the Naturejobs.com newsletter

and social media. Respondents hailed from a variety of scientific fields, but the basic biological sciences were most heavily represented.

Across all disciplines, PhD students became less pleased with their experience as their degrees progressed. Of first-year students who responded to the survey, 76% were "satisfied" or "very satisfied"; that decreased to 66.8% for second-years and 61.3% for third-years, although the numbers varied with region (see 'Continental divide'). Jo Rae Wright, dean of graduate-school programmes in cell biology at Duke University in Durham, North Carolina, was struck that the biggest drop was after the first year. "Students sometimes don't know what to expect", she says, and the reality can fail to live up to their aspirations. Hugh Kearns, a psychologist at Flinders University in Adelaide, Australia, who studies the graduate-student experience, says that the change could also be due to research results not turning out as expected. He notes that new students sometimes have unrealistically optimistic ideas about the feasibility of their research aims.

Satisfaction with "degree of independence" fluctuates. More than 72% of first-years were very or somewhat satisfied; that dipped to 64.1% for third-years, but rose again to 67.9% for fourth-years. What might be going on? "It's the nature of graduate education that stress and independence are linked," says Wright. She suggests that, in the United States anyway, as students get to the end of their graduate degrees, they focus more on their individual projects. This means more independence, but also a rise in anxiety: the pressure is on to conduct a successful project while sorting out future education, a postdoc or job prospects.

## BETTER ADVICE, BETTER OUTCOMES

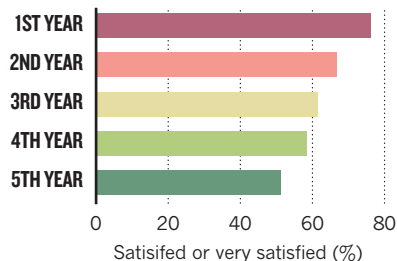
Advisers might not be doing enough to help graduate students to navigate the difficult later stages of their degrees — at least, the survey suggests that students perceive a deficiency. Among respondents in the first year of their PhDs, 62.6% were very or somewhat satisfied with the guidance that they received from their advisers; only 52.9% of second-years said the same, and 46.7% of third-years. The decline levelled off for fourth- and fifth-years, with about 43% very or somewhat satisfied. Likewise, satisfaction dipped noticeably when it came to "recognition from adviser". About 65% of first-years were somewhat or very satisfied, compared with only 48.6% of fourth-years.

Adviser recognition is an "essential element" of quality supervision, says Marja Makarow, ►

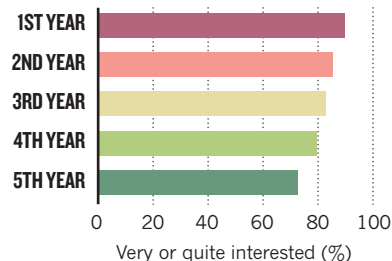
## Changing attitudes

Below are selected statistics from our international survey of graduate students. We had more than 5,000 respondents, with a heavy representation in the life sciences.

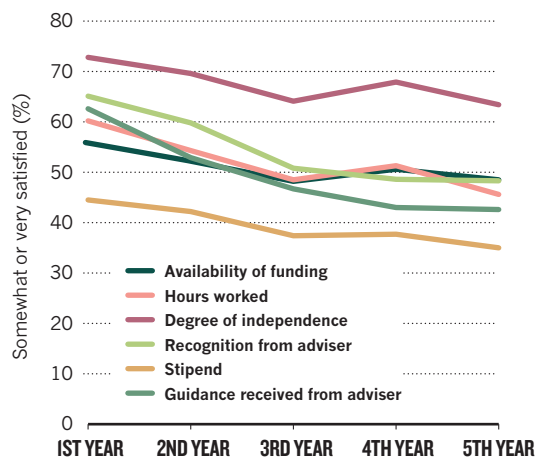
**Q** Overall, how satisfied are you with your graduate-school experience?



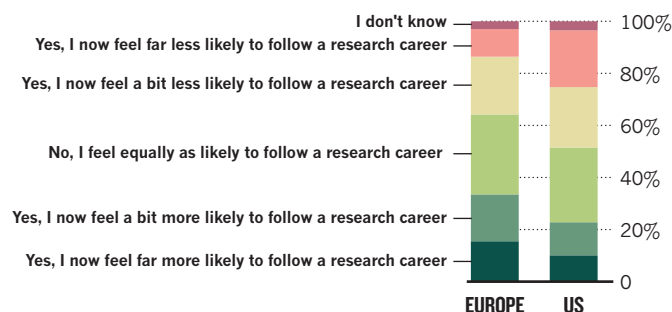
**Q** How would you rate your current interest in your work?



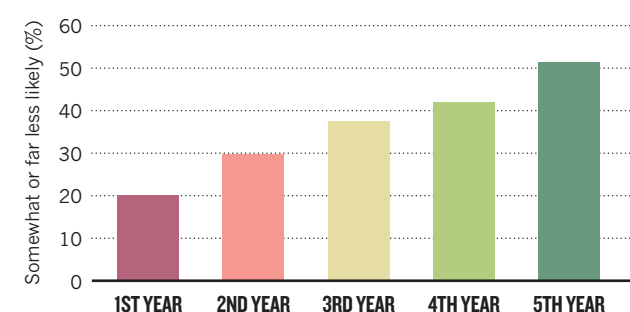
**Q** To what extent are you satisfied with each of the following attributes or aspects of your graduate-school experience?



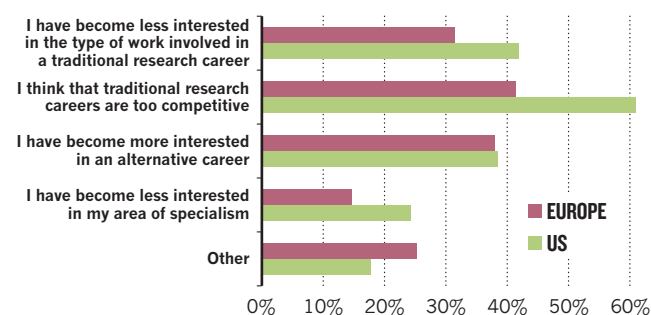
**Q** Do you think that your likelihood of commencing a research career has changed since starting your PhD?



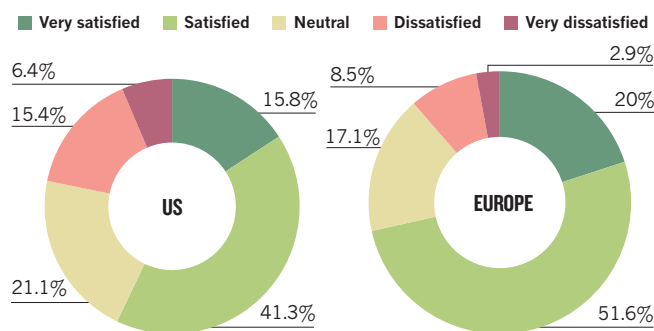
**Q** How much less likely to commence a research career have you become since starting your PhD?



**Q** Which, if any, of the following are reasons that you feel less likely to follow a traditional research career than you did when you started your PhD?



**Q** Overall, how satisfied are you with your graduate-school experience? By region.



► chief executive of the European Science Foundation in Strasbourg, France, and a long-time student adviser as professor of molecular biology at the University of Helsinki. "If there's no mutual valuing, you're missing something terribly important."

Kearns and his collaborator Maria Gardiner — with whom he runs workshops for graduate students and advisers and writes occasional columns for *Nature's* Careers section — say that the survey's findings are consistent with their research, which highlights how crucial a good adviser is to the graduate experience.

They have found that a lack of direction and clear advice from an adviser leads to significant declines in student satisfaction.

Kearns and Gardiner also find that many PhD students gradually lose confidence because they get little feedback about their performance. "This may, in turn, manifest itself in reduced satisfaction," says Kearns.

Advisers can minimize stress for their students by helping them to map out possible career aspirations. "We do need to do what we can to make that experience as rich and rigorous as possible, but also as humane for students as

possible," says Debra Stewart, president of the US Council of Graduate Schools in Washington DC. University deans and career centres should help faculty members to inform students about non-academic career tracks, she says. Faculty members often encourage students to follow the academic route not only because it is what they themselves love, but also because it is what they know best. "We have to help them paint a richer picture of careers," says Stewart.

Graduate students, for their part, need to consider their career options earlier, says Jennifer Rohn, a research fellow at the MRC Labo-



ratory for Molecular Cell Biology at University College London. Just 44.6% of respondents started thinking about their job preferences before they began their graduate degrees; 19.9% did so in their first year of study and 15.6% during their second. Too many left the decision until later in their degrees, says Rohn, who has advocated a shake-up of the postdoctoral career structure (see *Nature* 471, 7; 2011). “To me, this implies they are not entering [postgraduate education] to train for a trade — they are drifting into it because they don’t have alternative plans, or it’s just ‘the next step’ somehow and they can defer making a decision for a few more years,” she says. “Given how poor the job prospects are in academia, this is a concern.”

But to Wright, 44.6% is a surprisingly high number of students considering their future careers before they start their PhDs. “That’s not my experience,” she says. “Our students say they don’t consider until they near the finish line.” She points out that perhaps the new student has “thought about” career aspirations only fleetingly. They might not have followed up by actively researching options or discussing them with mentors and others in the know. Wright suspects that few students do so. She hopes to change that at Duke through a coordinated effort to expand career-development counselling, involving not only faculty members, but also alumni and the university careers centre.

Thomas Skalak, vice-president for research at the University of Virginia in Charlottesville, emphasizes the need to impress upon students that they are, in the end, responsible for their own education. He likes to suggest that they act as ‘intellectual entrepreneurs’ by fastidiously minding their own education, graduate project, research focus and career prospects.

### ACADEMIC BOOM AND BUST

Academic careers are still highly sought after, although graduate students are by no means ignoring industry and other options. About 57% said that they would pursue a postdoc in academia after graduating; 17.5% preferred a postdoc in industry. About 53% said that before starting their graduate education, they had wanted to end up in academia. Rohn wonders whether enough of these respondents had been properly briefed about the very strong competition that surrounds academic jobs. “By and large, graduate school is still perceived largely as a gateway to an academic career, and nobody is disabusing [students] of this notion,” she says. “Or if they are, people aren’t listening. It’s human nature to believe you can be the exception.”

The survey implies that the longer students spend in graduate education, the less attractive an academic career becomes. About 78.8% of first-year PhD students said they were “very” or “quite” likely to pursue a university research career; that dips to 70.3% for second-years. The

## REGIONAL DIFFERENCES

### Continental divide

Graduate students in Europe seem to have a more positive attitude towards their work than those in the United States, according to *Nature’s* graduate-student survey.

Of Europe-based students who responded, 71.6% were somewhat or very satisfied with their overall experience, compared with 57.1% in the United States and 62.3% in Japan. Whereas 18.2% of US-based respondents said that their experience had “significantly worsened” since they had started their degrees, just 8.3% of Europe-based respondents and 4.4% of respondents in Japan agreed. Graduate students in Europe also reported greater interest in their work — 46.1% were very interested, compared with only 32.7% in the United States.

When asked whether they had become less or more likely to continue on to a research career since they became graduate students, 21.8% in the United States replied that they were far less likely; just 10.6% of Europe-based respondents said so.

Differences in cultures, lifestyles and institutions make it difficult to pinpoint the

causes of the discrepancy. However, Marja Makarow, chief executive of the European Science Foundation in Strasbourg, France, suggests that a move towards more-structured PhD programmes in many parts of Europe in recent years could have boosted satisfaction there. At many European institutions, says Makarow, tutelage has expanded beyond the conventional one-on-one with an adviser, and now includes a group of faculty members. Some previously unstructured programmes now have workshops and lab courses. Also, getting a PhD typically takes three to four years in parts of Europe, whereas it can take five or more in the United States, which can cause dismay.

Far more students in the United States say that intense competition discourages them from pursuing a research career than do those in Europe. Among the US-based respondents who had become less likely to pursue a research career, 60.9% thought that “traditional careers are too competitive”, compared with just 41.3% of Europe-based respondents. **G.R.**

decline continues throughout the degree, with 66.8% of third-years, 64.5% of fourth-years and 62% of fifth-years answering the same.

Intense competition for original results, publications and jobs seems to be a major factor in this change. Among those respondents who had grown less likely to follow a conventional research career since starting their graduate degrees, 49.4% attributed their change of heart to competition, 36.9% had become more interested in an alternative career and 36.4% had become less interested in the work involved in traditional research, although there were significant regional discrepancies. (Respondents could choose more than one reason.)

### CAREER QUALMS

Data from the University of California, San Francisco, suggest other reasons for turning away from academia. In 2008, Cynthia Fuhrmann, the university’s programme director for academic career development, polled biomedical-research graduate students about their career preferences. As in *Nature’s* survey, she saw a drop in interest in academia. Among the 469 respondents, 42% of first-years wanted to be a “principal investigator at a research-intensive institution”; that dropped to 25% for third-year students. Of those who gave reasons, many cited the long work hours required, the challenge of getting funding, a distaste for daily tasks such as grant writing and the slow pace of research, and the intense competition for

tenure. Some also had what Fuhrmann terms “positive” reasons for their change of preference — such as learning about an exciting new job opportunity. The results will be published in September in *CBE — Life Sciences Education*.

Again, mentorship can make a difference — and have a big influence on career directions. Fuhrmann has seen more than one student who was doubtful about pursuing academia decide to go ahead after talking to a mentor. She suggests that more transparency from mentors — not only about successes, but also about disappointments — could help students resolve to stick with their plans. For example, when principal investigators do not get funded, they often shield their protégés from the failure. Perhaps they should instead share such undesirable outcomes with their trainees and explain how they plan to press forward, change tactics or find bridge funding. “Trainees notice these things anyway,” says Fuhrmann, “and it could be helpful for principal investigators to share how they deal with stressors.”

Around the world, students trying to complete a PhD in the sciences often face a daunting task. The survey data suggest that although they relish their independence, such students consistently crave wise counsel. Given tight university budgets and a scientific workforce and opportunities that continue to evolve, advice is as important now as it has ever been. ■

**Gene Russo** is *Careers* editor at *Nature*.

➔ **NATURE.COM**  
For more data tables  
from the graduate  
survey, visit:  
[go.nature.com/eqmi3e](http://go.nature.com/eqmi3e)

# TERMINAL TALK

*Getting to grips with history.*

BY SCOTT AKALIS

Snow fell on the adjacent hills, sleet on the airport. After they announced the delay, she smiled and asked: "Still drinking Coke?"

The motley beer-tap handles looked like a group of childhood friends grown apart. He turned away from them and responded: "Okay."

As they sipped their sodas, she studied his eyes — blue like hers. "Did you finish all your work?"

"Everything except this essay on abortion. It's just hard to write about something with no bearing on you."

She coughed up some cola. "You know, it was considered an intractable issue only 20 years ago."

"I guess."

"Do you have any choice?"

Staring out at the amorphous white hills, he shook his head. "What were the different sides, again?"

"I was worrying they taught you not to ask questions."

He rolled his eyes and pulled out his jEverything, a cheap iAll knockoff.

"Pro-life and pro-choice," she said.

"Or anti-choice and anti-life."

"Yeah. It's weird — we used to avoid talking about this, and your grandparents —"

"They didn't actually have one, did they?"

"No, they had *me*."

His eyes blinked and broadened.

"They were young and backpacking around Europe. Neither ever discussed it directly; that's just not how their generation was. But you could read..." She picked up her Coke and turned away, towards the overhead TV. One team played another.

He planted a hand against the bar and swivelled towards the concourse, his eyes fixing on a passing babe and the pink baggage trailing her.

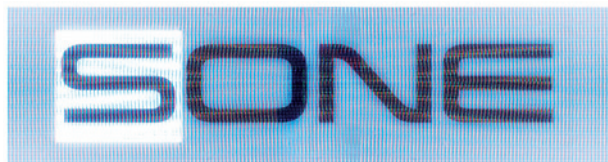
"Now it seems almost quaint, like that gold-standard thing," she said, turning back from the TV.

He reached for his Coke without rotating away from the concourse.

Her fragile smile broke. "Anyway, the pro-lifers believed life began at conception, which made abortion murder. For the pro-choice camp, it was more of a women's rights issue."

His straw choked on air.

"Have another," she said. "It's dry up there."



He sighed and swivelled back. "Was it, like, an age divide?"

"More a religious thing," she said with a slight lisp.

He checked his jEverything for transcription errors but found none.

"You have to understand that doctors were getting shot over this. Nobody expected a resolution, especially not without some consensus on when life begins."

He studied the cocktail garnishes below and the liquor bottles above. The wax seal on the Maker's Mark resembled a bleeding maraschino.

"Progress, if any, was supposed to come from scientific breakthroughs or shifts within the Church."

"Uh-huh."

"Are you listening to me?"

He deleted the last sentence and set the jEverything on the bar. She rose up and peered over. Its inferior screen limited the angles from which it could be read. Further thwarting her efforts, he slid his glass in front of the device and resumed their conversation. "Then came the sterility epidemic," he said.

The bartender, seeing the half-empty glass slide forward, rushed over and topped it up. She grinned and, after a moment, responded: "Right."

"Do you think someone engineered the virus?" he asked, lifting the Coke off the bar.

"Maybe, but I wouldn't cover the details of the conspiracy theories in your paper. I mean, half of them implicate that whacko who turned pregnant women's pee brown."

He set his drink down. She laughed.

"You haven't heard about this guy? He thought he could end late-term abortions by adding chemicals to his town's water supply that would turn every tinkle into a pregnancy test."

He looked dumbfounded. "So do I just skip the origins of the virus?"

"You could do a brief summary, but what's

going to get you an A is going beyond that — like asking whether the debate would have ended the way it did if

➔ **NATURE.COM**  
Follow Futures on  
Facebook at:  
[go.nature.com/mtoodm](https://go.nature.com/mtoodm)

God's fingerprints hadn't been on the epidemic."

He turned towards the window to contemplate. The outline of the crests appeared through the tapering snow. Closer, a worker finished spraying the underside of a wing.

"Anyway, you'll want to cover the epidemic itself: the onset and panic, the talk of extinction, the summits..." Her voice grew both more intimate and distant. "What a time to be alive."

"You mean right before I was born?"

"Yes. And never forget you have science to thank for that whole being-born thing."

"So I don't have to buy Mother's Day cards anymore?"

She grinned and shook her head. He finished his Coke and glanced towards the growing line of passengers at the gate.

"One more thing you should know is that *the pill* used to refer to a contraceptive, not a fertility treatment. I sometimes forget how your conception of pregnancy has always depended as much on those capsules as it has sperm."

His face reddened on eye contact with the bartender. He dismounted the stool and picked up his bag.

"Wait here a little longer," she said.

He set his bag on the stool.

"Now do you see the connections between this issue — obsolete as it may be — and its, uh, descendant?"

"Pill access?" he said. "Sure — some see it as a woman's right; others talk about the would-be baby and protest government subsidies, doctors who give the pill to minors —" "I think you're ready."

The intercom announced general boarding. As they hugged, he gazed out the window once more. The snow had stopped, and the glint of the ice was gone.

"Call me if you need more help," she said.

"Okay."

"Will you, please?"

"Yes." He took a step towards the gate and stopped. Turning around, he spotted his jEverything sitting on the bar.

She handed it to him, saying: "Wouldn't want to lose that one."

It transcribed 'one' as 'son'. ■

Scott Akalis's fiction has appeared in Camera Obscura, Underground Voices and Stymie Magazine. He has a PhD in psychology from Harvard University and lives in Chicago. [ScottAkalis.com](https://ScottAkalis.com).

JACEY

## CORRIGENDUM

doi:10.1038/nature10222

# Chromatin regulation by Brg1 underlies heart muscle development and disease

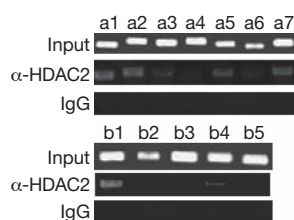
Calvin T. Hang, Jin Yang, Pei Han, Hsiu-Ling Cheng, Ching Shang, Euan Ashley, Bin Zhou & Ching-Pin Chang

*Nature* **466**, 62–67 (2010)

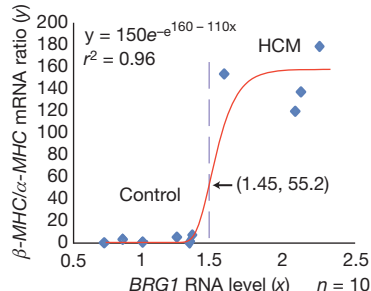
In this Letter, a panel in Fig. 4f (thymus chromatin immunoprecipitation (ChIP) input lane) was inadvertently duplicated in a panel in Fig. 4m (hearts with transverse aortic constriction (TAC) ChIP input lane). The corrected input lane for Fig. 4m, shown below, was from Supplementary Fig. 11f because the ChIPs for histone deacetylases 2 and 9 (HDAC2 and HDAC9) were conducted in the same experiment. Our practice is to establish an optimal polymerase chain reaction (PCR) condition for each sample so that all inputs attain a similar level of amplification regardless of tissue source, allowing normalization and cross-comparison among different samples. Because normalization involved only sonicated genomic DNA prior to immunoprecipitation, the bands from the input gels were nearly identical in appearance for both heart and thymus, and this unfortunately caused us to misplace the two input figures during manuscript preparation.

Also, in Fig. 5c the *y*-axis is mislabelled and should be ' $\beta$ -MHC/ $\alpha$ -MHC mRNA ratio (*y*)' and the mathematical formula should be  $y = 150e^{-e^{160-110x}}$  (the corrected Fig. 5c is shown below). These changes do not affect the overall conclusion of the manuscript. Nor do they affect our specific conclusions that the chromatin-remodelling protein Brg1 does not bind to myosin heavy chain (MHC)- $\alpha$  or MHC- $\beta$  promoters in thymus, whereas HDAC2 and HDAC9 bind to the same promoters in stressed adult hearts, and the mathematical model describing Brg1 and MHC expression.

**4m** Hearts with TAC



**5c**





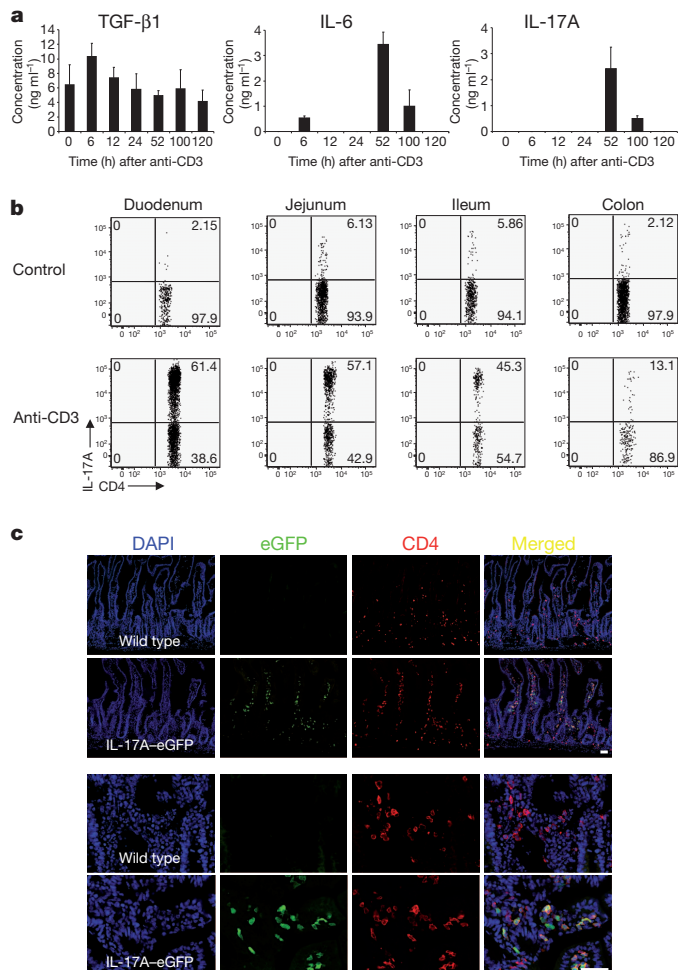
# Control of T<sub>H</sub>17 cells occurs in the small intestine

Enric Esplugues<sup>1,2,3\*</sup>, Samuel Huber<sup>1,4\*</sup>, Nicola Gagliani<sup>5</sup>, Anja E. Hauser<sup>2</sup>, Terrence Town<sup>6,7</sup>, Yisong Y. Wan<sup>8</sup>, William O'Connor Jr<sup>1</sup>, Anthony Rongvaux<sup>1</sup>, Nico Van Rooijen<sup>9</sup>, Ann M. Haberman<sup>10</sup>, Yoichiro Iwakura<sup>11</sup>, Vijay K. Kuchroo<sup>12</sup>, Jay K. Kolls<sup>13</sup>, Jeffrey A. Bluestone<sup>14</sup>, Kevan C. Herold<sup>1</sup> & Richard A. Flavell<sup>1,15</sup>

Interleukin (IL)-17-producing T helper cells (T<sub>H</sub>17) are a recently identified CD4<sup>+</sup> T cell subset distinct from T helper type 1 (T<sub>H</sub>1) and T helper type 2 (T<sub>H</sub>2) cells<sup>1</sup>. T<sub>H</sub>17 cells can drive antigen-specific autoimmune diseases and are considered the main population of pathogenic T cells driving experimental autoimmune encephalomyelitis (EAE)<sup>2</sup>, the mouse model for multiple sclerosis. The factors that are needed for the generation of T<sub>H</sub>17 cells have been well characterized<sup>3–6</sup>. However, where and how the immune system controls T<sub>H</sub>17 cells *in vivo* remains unclear. Here, by using a model of tolerance induced by CD3-specific antibody, a model of sepsis and influenza A viral infection (H1N1), we show that pro-inflammatory T<sub>H</sub>17 cells can be redirected to and controlled in the small intestine. T<sub>H</sub>17-specific IL-17A secretion induced expression of the chemokine CCL20 in the small intestine, facilitating the migration of these cells specifically to the small intestine via the CCR6/CCL20 axis. Moreover, we found that T<sub>H</sub>17 cells are controlled by two different mechanisms in the small intestine: first, they are eliminated via the intestinal lumen; second, pro-inflammatory T<sub>H</sub>17 cells simultaneously acquire a regulatory phenotype with *in vitro* and *in vivo* immune-suppressive properties (rT<sub>H</sub>17). These results identify mechanisms limiting T<sub>H</sub>17 cell pathogenicity and implicate the gastrointestinal tract as a site for control of T<sub>H</sub>17 cells.

T<sub>H</sub>17 cells have been associated with the pathogenesis of several chronic inflammatory disorders, including rheumatoid arthritis and multiple sclerosis<sup>2,7</sup>. To study the cellular and molecular mechanisms that control pathogenicity mediated by T<sub>H</sub>17 cells we first used the CD3-specific antibody treatment model. It is known that CD3-specific antibody treatment induces a 'cytokine storm' and local inflammation mainly in the small intestine<sup>8</sup>. Despite this it has been validated as an *in vivo* model of tolerization<sup>9</sup> and is now under study in human clinical trials<sup>10</sup>. By mimicking antigen, CD3-specific antibody treatment leads to activation-induced cell death (AICD) of T cells<sup>11,12</sup> and consequently a systemic upregulation of IL-6 (ref. 9) and transforming growth factor- $\beta$  (TGF- $\beta$ 1) induced by phagocyte engulfment of apoptotic T cells<sup>13</sup>. In line with these publications, we found that CD3-specific antibody treatment induced an immunoregulatory environment marked by simultaneous expression of TGF- $\beta$ 1 and IL-6 (Fig. 1a). The combination of these cytokines is important for the development of T<sub>H</sub>17 cells *in vitro* and *in vivo* as it has been previously clearly established<sup>3,4</sup>. Accordingly, we found elevated levels of IL-17A in plasma of CD3-specific antibody-treated animals compared to controls (Fig. 1a).

First, we aimed to investigate the source of IL-17A. It has been reported that a few hours after injection of CD3-specific antibody, there is a rapid disappearance of the majority of T cells from the



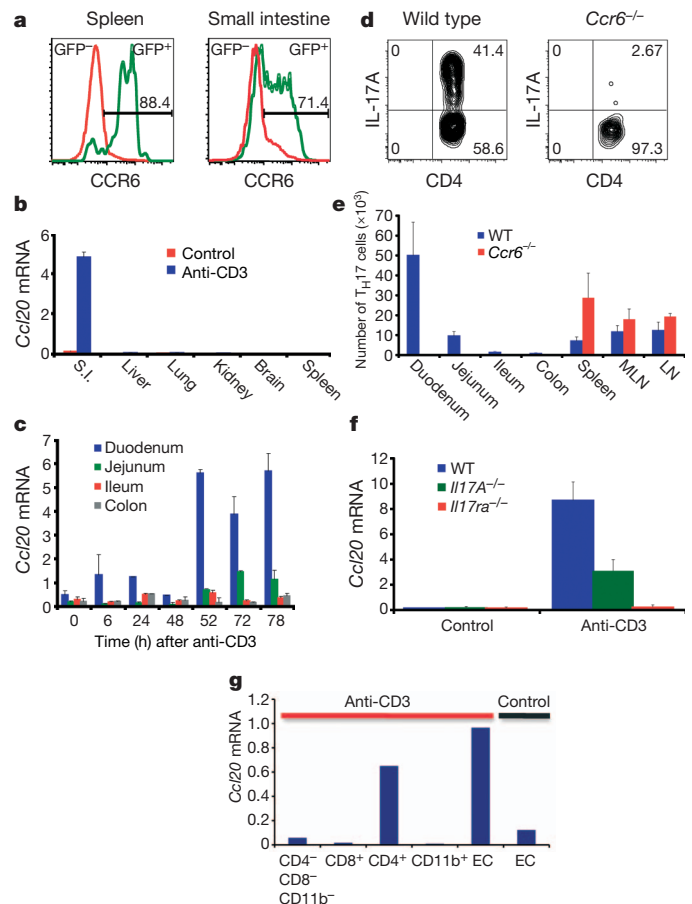
**Figure 1 | Accumulation of T<sub>H</sub>17 cells in the small intestine after CD3-specific antibody treatment.** Mice were injected with CD3-specific antibody. **a**, Plasma levels of TGF- $\beta$ 1, IL-6 and IL-17A. Mean  $\pm$  s.e.m.;  $n = 4$ . **b**, Flow cytometric analysis of IL-17A-eGFP expression (gated on CD4<sup>+</sup>TCR $\beta$ <sup>+</sup> events); numbers in quadrants indicate percent cells in each. **c**, Immunofluorescence staining of frozen sections of the small intestine after CD3-specific antibody treatment (eGFP, green; CD4, red; cell nuclei, DAPI). Scale bar, 50  $\mu$ m. Data are representative of at least three independent experiments.

<sup>1</sup>Department of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06520, USA. <sup>2</sup>German Rheumatism Research Center (DRFZ), A Leibniz Institute, Berlin 10117, Germany. <sup>3</sup>Cluster of Excellence NeuroCure, Charité-Universitätsmedizin Berlin, Berlin 10117, Germany. <sup>4</sup>Medizinische Klinik, Universitätsklinikum Hamburg-Eppendorf, Hamburg 20246, Germany. <sup>5</sup>San Raffaele Diabetes Research Institute (HSR-DRI), Milan 20132, Italy. <sup>6</sup>Departments of Biomedical Sciences and Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA. <sup>7</sup>Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, California 90048, USA. <sup>8</sup>Lineberger Comprehensive Cancer Center, Department of Microbiology and Immunology, The University of North Carolina, School of Medicine, Chapel Hill, North Carolina 27599, USA. <sup>9</sup>Department of Molecular Cell Biology, Faculty of Medicine, Vrije Universiteit, Amsterdam, 1081 BT, The Netherlands. <sup>10</sup>Department of Laboratory Medicine, Yale University School of Medicine, New Haven, Connecticut 06520, USA. <sup>11</sup>Center for Experimental Medicine, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan. <sup>12</sup>Center for Neurologic Diseases, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>13</sup>Department of Genetics, LSU Health Sciences Center, New Orleans, Louisiana 70112, USA. <sup>14</sup>Diabetes Center at the University of California San Francisco, San Francisco, California 94143, USA. <sup>15</sup>Howard Hughes Medical Institute, New Haven, \*These authors contributed equally to this work.

circulation<sup>13,14</sup>. Surprisingly, in parallel with the disappearance of T cells from the periphery, we found a concomitant increase in the percentage and the number of total T cells in the small intestine, in particular in the duodenum (Supplementary Fig. 1a–c). In a newly generated IL-17A–eGFP knock-in mouse (enhanced green fluorescent protein was inserted in the *Il17a* locus; Methods and Supplementary Fig. 2a–d and 3a–c) injected with CD3-specific antibody, 50–80% of the CD4<sup>+</sup>TCRβ<sup>+</sup> T cells located in the duodenum were expressing IL-17A (Fig. 1b and Supplementary Fig. 1d, e). The percentage and number of T<sub>H</sub>17 cells in the intestine decreased from the duodenum to the colon in a gradient-like fashion (Fig. 1b). Detection of CD4<sup>+</sup>eGFP<sup>+</sup> T cells by immunofluorescence and two-photon-laser-scanning microscopy confirmed the high frequency of T<sub>H</sub>17 cells in the small intestine *in situ* (Fig. 1c and Supplementary Fig. 4a–c). Importantly, we also found T<sub>H</sub>17 cell infiltration in the duodenum when animals were injected with a therapeutic non-FcR-binding CD3-specific antibody<sup>15</sup>, although the frequency and numbers of the T<sub>H</sub>17 cells were lower compared to the FcR-binding antibody (Supplementary Fig. 5a). Similar results were observed after antigen-specific stimulation when soluble myelin oligodendrocyte glycoprotein antigen (MOG) was administered to MOG-TCR transgenic mice (2D2 mice)<sup>16</sup> (Supplementary Fig. 5b). Taken together these data suggest that the generation and the accumulation of T<sub>H</sub>17 cells in the small intestine was not restricted to the CD3-specific antibody treatment, but was a general mechanism following strong T-cell receptor (TCR) stimulation.

We next wanted to identify the molecular signals important for the generation of T<sub>H</sub>17 cells *in vivo* after CD3-specific antibody treatment. Because IL-6 is known to be important for T<sub>H</sub>17 cell generation, we evaluated the importance of this cytokine. *Il6*<sup>−/−</sup> and wild-type mice were treated with CD3-specific antibody. In the *Il6*<sup>−/−</sup> mice, only a very small population of T<sub>H</sub>17 cells (about 2%) could be found by flow cytometry in the small intestine (Supplementary Fig. 6a) and IL-17A was undetectable in the plasma (data not shown). To study the cellular source of IL-6, we treated mice with clodronate-loaded liposomes, which eliminates most macrophages and a significant proportion of dendritic cells compared to PBS-loaded liposomes<sup>13</sup> (Supplementary Fig. 6c). IL-6 plasma levels were greatly reduced in mice treated with clodronate-loaded liposomes compared to control mice after CD3-specific antibody injection (Supplementary Fig. 6d) and a profound reduction in T<sub>H</sub>17 cells was observed (Supplementary Fig. 6b, c). Taken together, these data support the notion that IL-6 secreted by antigen-presenting cells (APCs) is critical for the generation of T<sub>H</sub>17 cells during CD3-specific antibody treatment.

We next analysed the mechanism leading to the specific accumulation of T<sub>H</sub>17 cells in the small intestine, predominantly in the duodenum. T<sub>H</sub>17 cells are known to express the chemokine receptor CCR6 (ref. 17). Whereas CCR6 is relevant in different autoimmune disease models<sup>7,18</sup>, the role of the CCR6/CCL20 axis in immune cell migration to the intestine during tolerance induction has not yet been evaluated. To study that, we analysed the expression of CCR6 on CD4<sup>+</sup> IL-17A–eGFP positive and negative cells (Fig. 2a) and *Ccl20* mRNA expression (Fig. 2b) in the spleen and the gut. CCR6 was mainly expressed in T<sub>H</sub>17 cells from the spleen and the gut 24 h after CD3-specific antibody injection (Fig. 2a). Strikingly, when we performed a time course to measure the mRNA levels of *Ccl20* in different parts of the intestine during CD3-specific antibody treatment, we observed that *Ccl20* was expressed at the highest level in the duodenum in steady state conditions and was selectively further upregulated after CD3-specific antibody treatment (Fig. 2b, c and Supplementary Fig. 7). To test the importance of the CCR6/CCL20 axis for the migration of T<sub>H</sub>17 cells from the periphery to the duodenum, we treated *Ccr6*<sup>−/−</sup> and control mice with CD3-specific antibody. T<sub>H</sub>17 cell number (Fig. 2e) and frequency (Fig. 2d) were strongly reduced in the intestine of the *Ccr6*<sup>−/−</sup> compared to wild-type mice. In general, we did not observe signs of intestinal inflammation in the *Ccr6*<sup>−/−</sup> mice as we did



**Figure 2 | The axis CCR6/CCL20 is essential for the recruitment of T<sub>H</sub>17 cells to the small intestine.** **a**, CCR6 expression 24 h after anti-CD3 treatment. **b**, **c**, *Ccl20* mRNA expression (mean ± s.e.m.; *n* = 4). S.I., small intestine. **d**, IL-17A expression (gated on CD4<sup>+</sup>TCRβ<sup>+</sup> events) as measured by intracellular cytokine staining. **e**, T<sub>H</sub>17 cell numbers in different organs (mean ± s.d.; *n* = 5). LN, lymph node; MLN, mesenteric lymph node. **f**, *Ccl20* mRNA expression in duodenum of wild-type (WT), *Il17a*<sup>−/−</sup> and *Il17ra*<sup>−/−</sup> mice (mean ± s.e.m.; *n* = 4). **g**, *Ccl20* mRNA levels of epithelial and haematopoietic cells isolated from the small intestine. EC, epithelial cells. Panels **b**, **d**–**g** show results 100 h after the first anti-CD3 injection. Data are representative of at least three independent experiments.

in wild-type controls after CD3-specific antibody treatment (data not shown). Interestingly, we detected a higher number of T<sub>H</sub>17 cells in the spleen and lymph nodes of *Ccr6*<sup>−/−</sup> mice when compared to control animals (Fig. 2e). This increase was accompanied by splenomegaly and enlargement of lymph nodes (data not shown), indicating that CCR6 does not have a major role in the generation and expansion of T<sub>H</sub>17 cells. In conclusion, CCR6 seems to be essential for the migration of T<sub>H</sub>17 cells to the small intestine after CD3-specific antibody treatment, and the intestinal inflammation is dependent on this migration. Thus our data indicate that T<sub>H</sub>17 cell migrate to the small intestine leading to intestinal inflammation and damage. However, we cannot exclude that a proliferation of gut resident T<sub>H</sub>17 cells also contributes to the observed phenomenon.

To evaluate the contribution of IL-17A and IL-17F (T<sub>H</sub>17 signature cytokines) in the induction of CCL20 expression in the duodenum, we treated *Il17a*<sup>−/−</sup> or *Il17ra*<sup>−/−</sup> mice with CD3-specific antibody. We found decreased levels of CCL20 in the *Il17a*<sup>−/−</sup> and the *Il17ra*<sup>−/−</sup> mice versus the controls after CD3-specific antibody treatment (Fig. 2f), indicating that IL-17 signalling has a major role in the induction of CCL20 in the duodenum. We next studied the cellular source of CCL20. *Ccl20* mRNA was only detectable in the intestinal epithelial cells in untreated mice. Treatment with CD3-specific antibody led to a

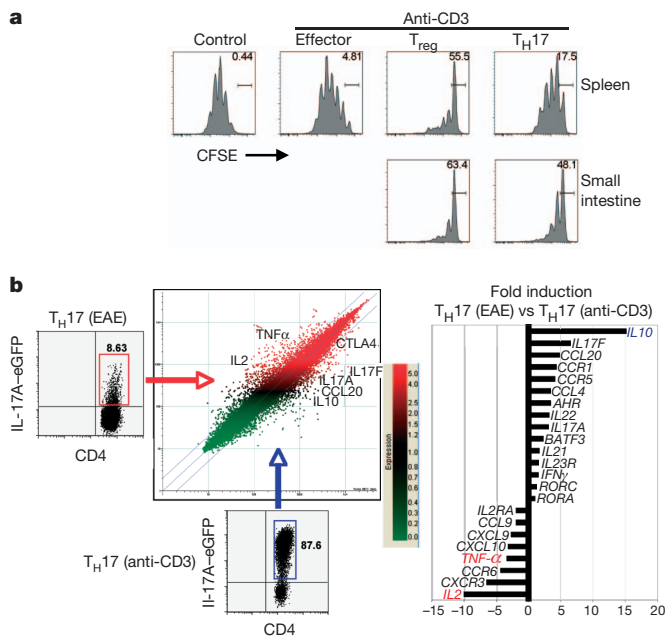
further upregulation of *Ccl20* mRNA by the epithelial cells. Additionally, the CD4<sup>+</sup> T cells present in the small intestine after CD3-specific antibody treatment, most of which were T<sub>H</sub>17 cells, expressed high levels of *Ccl20* mRNA (Fig. 2g). In conclusion, T<sub>H</sub>17 cells via IL-17A and IL-17F production directly upregulate CCL20 production by the intestinal epithelial cells, which then leads to the subsequent recruitment of CCR6<sup>+</sup> T<sub>H</sub>17 cells, which also produce CCL20.

Of note, the intestinal inflammation after CD3-specific antibody treatment was transient and 100% of the mice recovered. To understand better the mechanisms underlying this process, we first assessed apoptosis of T<sub>H</sub>17 cells in the small intestine but we did not detect a significant number of apoptotic cells (data not shown). When we studied the *in vivo* proliferation capacity of CD4<sup>+</sup>TCRβ<sup>+</sup> T cells from the CD3-specific antibody-treated animals, we found that T<sub>H</sub>17 cells from the duodenum were actively proliferating (Supplementary Figs 8 and 9a, b). Using IL17A-eGFP × FoxP3-mRFP double reporter mice (monomeric red fluorescent protein was inserted in the *foxp3* locus) we determined that CD4<sup>+</sup>IL-17A<sup>+</sup> T cells were proliferating at a higher rate than CD4<sup>+</sup>IL-17A<sup>-</sup> T cells in the duodenum (Supplementary Figs 8 and 9). Using two-photon laser-scanning microscopy, we found that the T<sub>H</sub>17 cells in the duodenum did not show the typical behaviour of an apoptotic T cell, conversely, they behaved like activated T cells in terms of their pattern of speed and direction of migration (Supplementary Video). Taken together these data indicate that T<sub>H</sub>17 cells do not die in the small intestine, but are rather actively proliferating.

In line with previous publications<sup>8</sup> we found that CD3-specific antibody treatment caused diarrhoea, oedema, inflammation and tissue destruction in the small intestine (Supplementary Fig. 10a, b), which correlated with the recruitment of T<sub>H</sub>17 cells. However, the intestinal pathology was only transient and mice fully recovered. We therefore began to investigate the fate of T<sub>H</sub>17 cells in the small intestine. Interestingly, we found a fraction of T<sub>H</sub>17 cells in the intestinal lumen

of the CD3-specific antibody-treated mice (Supplementary Fig. 10c, d). Given the severe inflammation, diarrhoea and tissue damage (Supplementary Fig. 10a, b), it is most likely that these cells were passively washed out, although an active mechanism cannot be excluded. Considering that the remaining T<sub>H</sub>17 cells in the duodenum were actively proliferating, but the intestinal pathology was only transient, we were curious about the functional capabilities of these cells. Surprisingly, we found that the remaining T<sub>H</sub>17 cells in the duodenum were able to suppress proliferation of responder T cells *in vitro* (Fig. 3a). To study the molecular properties of these suppressive T<sub>H</sub>17 cells (which we refer from now on as rT<sub>H</sub>17 cells), we performed a genome-wide transcriptional profiling assay (Fig. 3b). We compared the gene expression pattern of rT<sub>H</sub>17 cells from CD3-specific antibody-treated mice and genes expressed by pro-inflammatory T<sub>H</sub>17 cells that were harvested from the central nervous system of EAE-induced mice. The signature genes of T<sub>H</sub>17 cells, like *Rorc*, *Rora*, *Il17a*, *Il22* or *Il23r*, were similarly expressed between both types of T<sub>H</sub>17 cells. Also the activation status of these cells seemed to be similar, because activation markers such as CD69, CD25 and CD44 were equally expressed. However, we found that the rT<sub>H</sub>17 cells from the CD3-specific antibody-treated mice showed a non-inflammatory gene expression profile compared to pro-inflammatory T<sub>H</sub>17 cells isolated from the central nervous system of EAE-induced mice. Notably, the expression levels of *Tnf-α* and *Il-2*, two cytokines with clear pro-inflammatory roles<sup>19,20</sup>, were greatly reduced in the rT<sub>H</sub>17 cells from the small intestine. In contrast, these cells expressed high levels of IL-10, a cytokine with potent anti-inflammatory activities<sup>21</sup> (Supplementary Fig. 11b and Fig. 3b). These data are supported by a previous report showing that *in-vitro*-generated non-pathogenic T<sub>H</sub>17 cells are able to express IL-10 (ref. 22). To evaluate the molecular mechanisms involved in the suppressive function of the rT<sub>H</sub>17 cells, different molecules were blocked in an *in vitro* suppression assay using monoclonal antibodies (Supplementary Fig. 11a). The suppressive capacity of the rT<sub>H</sub>17 cells was partially dependent on IL-10, CTLA-4 and TGF-β. Blocking all three pathways resulted in a lack of suppression by the rT<sub>H</sub>17 cells. T<sub>H</sub>17 cells isolated from the spleen showed an intermediate phenotype. They exhibited a limited capacity to suppress the proliferation of T cells *in vitro* (Fig. 3a), and also produced more TNF-α and IL-2, but less IL-10 compared to T<sub>H</sub>17 cells isolated from the small intestine (Supplementary Fig. 11b). However, because some of the T<sub>H</sub>17 cells in the small intestine downregulated CCR6 (Fig. 2a), it is possible that some rT<sub>H</sub>17 might have migrated back from the small intestine to the spleen. If development of the suppressive capability occurred in the small intestine, then preventing the migration of the T<sub>H</sub>17 cells to that site should prevent the development of these tolerogenic cells. To test this hypothesis, we analysed T<sub>H</sub>17 cells isolated from the spleen of *Ccr6*<sup>-/-</sup> mice, because we showed already that these T<sub>H</sub>17 cells are unable to migrate to the small intestine. Consistent with the hypothesis, *Ccr6*<sup>-/-</sup> T<sub>H</sub>17 cells in the spleen showed high TNF-α production, failed to suppress T-cell proliferation *in vitro*, and were even proinflammatory, causing inflammatory bowel disease *in vivo* upon transfer into a lymphopenic host (Supplementary Fig. 12 a–d). These data indicate that proinflammatory T<sub>H</sub>17 cells do indeed acquire their suppressive phenotype in the small intestine.

To confirm our findings in an animal disease model, EAE-induced mice were treated with CD3-specific antibody. In line with a previous publication<sup>23</sup> we observed a protective effect when the treatment was administered during the course of the disease (Supplementary Fig. 13a). More importantly, we demonstrated that T<sub>H</sub>17 cells were recruited to the duodenum of the CD3-specific antibody treated animals and these mice had strongly reduced numbers of T<sub>H</sub>17 cells in the central nervous system (data not shown). Using a MOG-specific tetramer, we determined that a significant percentage of T<sub>H</sub>17 cells in the duodenum were antigen-specific (Supplementary Fig. 13b), demonstrating that MOG-specific T<sub>H</sub>17 cells were recruited to the duodenum following CD3-specific antibody treatment. In contrast the frequency



**Figure 3 | Functional and molecular characterization of rT<sub>H</sub>17 cells.** **a**, Suppression assay was performed using eGFP<sup>+</sup>mRFP<sup>+</sup>CD4<sup>+</sup> (Effector), eGFP<sup>+</sup>mRFP<sup>+</sup>CD4<sup>+</sup> (Tregs) or eGFP<sup>+</sup>mRFP<sup>+</sup>CD4<sup>+</sup> (T<sub>H</sub>17) cells sorted from spleen or small intestine. (Bar represents undivided CFSE-labelled CD4<sup>+</sup>CD25<sup>+</sup> responder T cells). Data are representative of six independent experiments. **b**, Gene expression analysis comparing T<sub>H</sub>17 cells (eGFP<sup>+</sup>mRFP<sup>+</sup>CD4<sup>+</sup>) from central nervous system at day 17 after EAE induction versus T<sub>H</sub>17 cells isolated from the small intestine of anti-CD3 treated IL-17A-eGFP × Foxp3-mRFP double reporter mice.



of MOG-tetramer-positive  $T_H17$  cells was much lower in other organs of EAE-induced mice, which had not been treated with CD3-specific antibody (data not shown). This is evidence against a general increase in MOG-specific  $T_H17$ . Therefore our results show that antigen-specific  $T_H17$  cells, with proinflammatory properties, generated in the periphery can be redirected to the small intestine. To confirm that  $rT_H17$  cells isolated from the small intestine of CD3-specific antibody-treated mice are indeed *in vivo* immune-suppressive we tested their suppressive capacity in an EAE transfer model. We co-transferred MOG-specific *in-vitro*-differentiated  $T_H17$  either alone or together with MOG-specific  $rT_H17$  cells isolated from the small intestine of CD3-specific antibody-treated 2D2 transgenic mice. Strikingly, we found that  $rT_H17$  cells were able to completely suppress the development of EAE in these transfer experiments (Supplementary Fig. 13c, d), indicating that the  $rT_H17$  cells are indeed stable in terms of their immune suppressive function.

As mentioned above CD3-specific antibody treatment is already used in clinical trials<sup>9,10</sup>, and we therefore aimed to confirm our results using teplizumab (hOKT3 $\gamma$ 1(Ala-Ala)), one CD3-specific antibody used in these trials. To that end we used a humanized mouse system: we reconstituted Balb/c Rag-2<sup>-/-</sup> $\gamma$ c<sup>-/-</sup> double knockout mice with human peripheral blood mononuclear cells. Two weeks after the transfer we treated these mice with either OKT-3, an FcR-binding

CD3-specific antibody used in the first human studies, or teplizumab, an FcR non-binding CD3-specific antibody. Strikingly, we found human T cells in the small intestine after treatment with both of these CD3-specific antibodies (Supplementary Fig. 14a, b). The presence of human IL-17A-, IL-10- and CCL20-producing cells in the small intestine in OKT-3- and teplizumab-treated mice was confirmed by real-time PCR (Supplementary Fig. 14c).

Taken together our results obtained in the CD3-specific antibody model suggest that  $T_H17$  cells, by upregulating CCL20 expression in the duodenum via IL-17 signalling, have developed an elegant mechanism to limit the pathogenicity in order to avoid a life-threatening immune response. This predicted in turn that this mechanism should be general to most strong immune responses that result in  $T_H17$  cells.

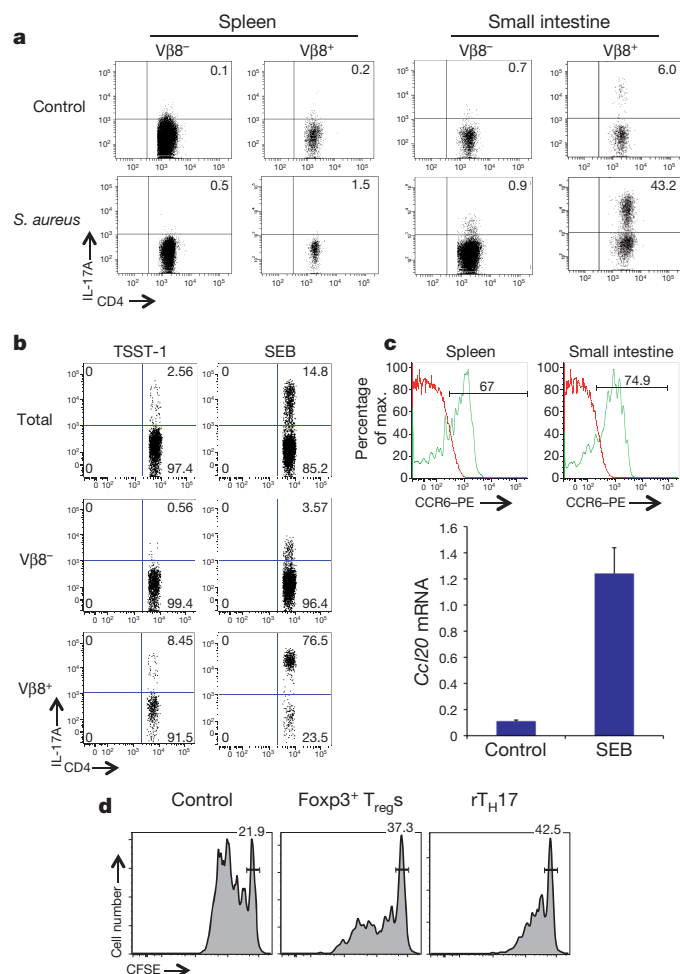
$T_H17$  cells have a crucial role in controlling different microorganisms *in vivo*<sup>24</sup>. We next investigated whether this mechanism of  $T_H17$  cell control also functions during a strong immune response elicited by pathogenic microorganisms. We first used a murine model of sepsis. We injected *Staphylococcus aureus*, which is one of the most frequent organisms responsible for sepsis in humans<sup>25,26</sup>, intravenously into IL-17A-eGFP reporter mice. Mice were analysed 3 days after the injection, at a time when they displayed severe clinical symptoms of sepsis (weight loss, dehydration, lethargy). Strikingly, we found the highest frequency and number of  $T_H17$  in the small intestine (Fig. 4a). Interestingly, most  $T_H17$  appeared to be TCR V $\beta$ 8<sup>+</sup>. The injection of the superantigen SEB (*Staphylococcus aureus* enterotoxin B), which is produced by the bacteria used in these experiments and binds to V $\beta$ 8<sup>+</sup> T cells, was sufficient to induce the accumulation of  $T_H17$  in the small intestine just as in the anti-CD3 studies. As a control we injected mice with TSST-1 (toxic shock syndrome toxin 1), a superantigen that does not bind to V $\beta$ 8 and is not produced by the bacteria we used. Of note, we observed that the administration of TSST-1 was less effective at inducing the accumulation of  $T_H17$  cells in the small intestine (Fig. 4b). Finally, we could confirm that the  $T_H17$  cells, induced by SEB treatment, expressed CCR6 and that CCL20 is specifically upregulated in the small intestine following SEB treatment (Fig. 4c). Furthermore, while a subpopulation of the  $T_H17$  cells was found in the intestinal lumen (Supplementary Fig. 15), the remaining  $T_H17$  cells demonstrated an immune-suppressive phenotype (Fig. 4d and Supplementary Fig. 16), again comparable to our results obtained in the CD3-specific antibody treatment. Interestingly, it is known that SEB can induce tolerance<sup>27</sup>, which is in line with our results that SEB leads to the generation of  $rT_H17$  cells. Accordingly, we found that SEB and, to a lesser extent, TSST-1 treatment of EAE-induced mice led to the amelioration of disease (data not shown), which is in line with one previous publication<sup>28</sup>.

In addition to anti-bacterial immunity, viruses are the next key class of pathogens to which we must respond, yet contain excessive immunopathology which is commonly the cause of morbidity and mortality<sup>29</sup>. To address such an immune response, we analysed influenza, a viral infection that has devastated human populations. Notably, we again found increased  $T_H17$  cell frequencies in the small intestine in mice infected with influenza A (H1N1) (Supplementary Fig. 17).

In conclusion, we propose a general mechanism that could explain how a pro-inflammatory  $T_H17$  immune response, which is beneficial in clearing infection, but immunopathogenic in excess, can be controlled by the mechanisms we describe here: namely by acquisition of an immune-suppressive phenotype or elimination into the intestinal lumen (Supplementary Fig. 18). These findings and further studies aiming to identify the underlying mechanism of the conversion of pro-inflammatory  $T_H17$  cells into  $rT_H17$  cells may help in designing new strategies to control auto-reactive  $T_H17$  cells in autoimmune diseases like multiple sclerosis.

## METHODS SUMMARY

**Anti-CD3, SEB, TSST-1 treatment and *Staphylococcus aureus* infection.** Mice were injected intraperitoneally three times with either CD3-specific antibody



**Figure 4 |  $T_H17$  cells are recruited to the small intestine during sepsis.** **a, b,** IL-17A-eGFP expression is shown (gated on CD4<sup>+</sup>TCR $\beta$ <sup>+</sup> events). Mice were injected with *Staphylococcus aureus* (**a**) or SEB and TSST-1 (**b**). **c,** CCR6 expression 24 h after the first SEB injection (top). *Ccl20* mRNA levels in the small intestine 100 h after the first injection (mean  $\pm$  s.e.m.;  $n = 4$ ) (bottom). **d,** *In vitro* suppression assay using CD4<sup>+</sup>IL-17A-eGFP<sup>+</sup> cells from the small intestine or CD4<sup>+</sup>Foxp3-mRFP<sup>+</sup> cells from the spleen of SEB-treated mice as suppressor cells. Results are representative of at least two independent experiments.

(clone 2C11, 20 µg per mouse) SEB (50 µg per mouse) or TSST-1 (50 µg per mouse) at 0, 48 and 96 h. Mice were analysed 100 h after the first injection, if not otherwise specified. *Staphylococcus aureus* was injected intravenously ( $1 \times 10^8$  colony-forming units per mouse) in order to induce sepsis. Mice were killed 3 days after the injection.

**Flow cytometric analysis.** Cells were isolated from the organ as indicated. IL-17A-eGFP and CCR6 expression was assessed directly after isolation. When indicated cells were restimulated and intracellular cytokine staining for IL-17A was performed. Numbers in dot-plot quadrants indicate percent cells in each. Cells were gated on CD4<sup>+</sup>TCRβ<sup>+</sup> events.

**Real-time PCR.** *Ccl20* mRNA expression was measured in different tissues as indicated using real time PCR with reverse transcription.

**In vitro suppression assay.** Different suppressor cells were co-cultured with carboxyfluorescein diacetate succinimidyl ester (CFSE)-labelled CD4<sup>+</sup>CD25<sup>-</sup> responder T cells, which were isolated from the spleen of CD45.1 congenic mice. Bar represents undivided CFSE-labelled responder T cells.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 12 May 2010; accepted 19 May 2011.

Published online 17 July 2011.

- Korn, T., Bettelli, E., Oukka, M. & Kuchroo, V. K. IL-17 and Th17 cells. *Annu. Rev. Immunol.* **27**, 485–517 (2009).
- Langrish, C. L. *et al.* IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. *J. Exp. Med.* **201**, 233–240 (2005).
- Bettelli, E. *et al.* Reciprocal developmental pathways for the generation of pathogenic effector Th17 and regulatory T cells. *Nature* **441**, 235–238 (2006).
- Mangan, P. R. *et al.* Transforming growth factor-β induces development of the Th17 lineage. *Nature* **441**, 231–234 (2006).
- Manel, N., Unutmaz, D. & Littman, D. R. The differentiation of human Th17 cells requires transforming growth factor-β and induction of the nuclear receptor RORγt. *Nature Immunol.* **9**, 641–649 (2008).
- Yang, L. *et al.* IL-21 and TGF-β are required for differentiation of human Th17 cells. *Nature* **454**, 350–352 (2008).
- Hirota, K. *et al.* Preferential recruitment of CCR6-expressing Th17 cells to inflamed joints via CCL20 in rheumatoid arthritis and its animal model. *J. Exp. Med.* **204**, 2803–2812 (2007).
- Merger, M. *et al.* Defining the roles of perforin, Fas/FasL, and tumour necrosis factor α in T cell induced mucosal damage in the mouse intestine. *Gut* **51**, 155–163 (2002).
- Chatenoud, L. & Bluestone, J. A. CD3-specific antibodies: a portal to the treatment of autoimmunity. *Nature Rev. Immunol.* **7**, 622–632 (2007).
- Herold, K. C. *et al.* Treatment of patients with new onset type 1 diabetes with a single course of anti-CD3 mAb teplizumab preserves insulin production for up to 5 years. *Clin. Immunol.* **132**, 166–173 (2009).
- Carpenter, P. A. *et al.* Non-Fc receptor-binding humanized anti-CD3 antibodies induce apoptosis of activated human T cells. *J. Immunol.* **165**, 6205–6213 (2000).
- Smith, C. A., Williams, G. T., Kingston, R., Jenkinson, E. J. & Owen, J. J. Antibodies to CD3/T-cell receptor complex induce death by apoptosis in immature T cells in thymic cultures. *Nature* **337**, 181–184 (1989).
- Perruche, S. *et al.* CD3-specific antibody-induced immune tolerance involves transforming growth factor-β from phagocytes digesting apoptotic T cells. *Nature Med.* **14**, 528–535 (2008).
- Chatenoud, L., Primo, J. & Bach, J. F. CD3 antibody-induced dominant self tolerance in overtly diabetic NOD mice. *J. Immunol.* **158**, 2947–2954 (1997).
- Alegre, M. L. *et al.* An anti-murine CD3 monoclonal antibody with a low affinity for Fc gamma receptors suppresses transplantation responses while minimizing acute toxicity and immunogenicity. *J. Immunol.* **155**, 1544–1555 (1995).
- Bettelli, E. *et al.* Myelin oligodendrocyte glycoprotein-specific T cell receptor transgenic mice develop spontaneous autoimmune optic neuritis. *J. Exp. Med.* **197**, 1073–1081 (2003).
- Annunziato, F. *et al.* Phenotypic and functional features of human Th17 cells. *J. Exp. Med.* **204**, 1849–1861 (2007).
- Reboldi, A. *et al.* C-C chemokine receptor 6-regulated entry of Th17 cells into the CNS through the choroid plexus is required for the initiation of EAE. *Nature Immunol.* **10**, 514–523 (2009).
- Brown, S. J. & Mayer, L. The immune response in inflammatory bowel disease. *Am. J. Gastroenterol.* **102**, 2058–2069 (2007).
- Petitto, J. M., Streit, W. J., Huang, Z., Butfiloski, E. & Schifflbauer, J. Interleukin-2 gene deletion produces a robust reduction in susceptibility to experimental autoimmune encephalomyelitis in C57BL/6 mice. *Neurosci. Lett.* **285**, 66–70 (2000).
- Pestka, S. *et al.* Interleukin-10 and related cytokines and receptors. *Annu. Rev. Immunol.* **22**, 929–979 (2004).
- McGeachy, M. J. *et al.* TGF-β and IL-6 drive the production of IL-17 and IL-10 by T cells and restrain Th17 cell-mediated pathology. *Nature Immunol.* **8**, 1390–1397 (2007).
- Kohm, A. P. *et al.* Treatment with nonmitogenic anti-CD3 monoclonal antibody induces CD4<sup>+</sup> T cell unresponsiveness and functional reversal of established experimental autoimmune encephalomyelitis. *J. Immunol.* **174**, 4525–4534 (2005).
- Khader, S. A., Gaffen, S. L. & Kolls, J. K. Th17 cells at the crossroads of innate and adaptive immunity against infectious diseases at the mucosa. *Mucosal Immunol.* **2**, 403–411 (2009).
- Rodriguez-Creixems, M. *et al.* Bloodstream infections: evolution and trends in the microbiology workload, incidence, and etiology, 1985–2006. *Medicine (Baltimore)* **87**, 234–249 (2008).
- Martin, G. S., Mannino, D. M., Eaton, S. & Moss, M. The epidemiology of sepsis in the United States from 1979 through 2000. *N. Engl. J. Med.* **348**, 1546–1554 (2003).
- Ochi, A., Yuh, K. & Migita, K. Not every superantigen induces tolerance *in vivo*. *Semin. Immunol.* **5**, 57–63 (1993).
- Soos, J. M., Schifflbauer, J. & Johnson, H. M. Treatment of PL/J mice with the superantigen, staphylococcal enterotoxin B, prevents development of experimental allergic encephalomyelitis. *J. Neuroimmunol.* **43**, 39–43 (1993).
- Chowell, G. *et al.* Severe respiratory disease concurrent with the circulation of H1N1 influenza. *N. Engl. J. Med.* **361**, 674–679 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** The authors would like to thank F. Manzo for expert administrative assistance, L. Evangelisti and C. Hughes for generating embryonic stem cells and chimaeric mice, respectively, J. Stein for initial screening of knock-in mice, T. Ferrandino for assistance with the mouse colony, E. Eynon and J. Alderman for managing the mouse program and A. Lin for assistance with Gene Array analysis. We also thank T. Taylor and G. Tokmouline for expert help with the FACS sorting and D. Gonzalez for help with the multiphoton microscopy. We would like to thank J. P. Allison for providing the anti-CTLA-4 antibody, F. Waldron-Lynch and J. S. Pober for providing peripheral blood mononuclear cells, and the NIH Tetramer core facility for providing the tetramers. W.O. was supported by a fellowship from the National Multiple Sclerosis Society. S.H. was supported by the DFG (HU 1714/1-1) and by a James Hudson Brown-Alexander B. Coxie Fellowship. E.E. was supported by the Spanish Ministry of Science postdoctoral fellowship and by a James Hudson Brown-Alexander B. Coxie Fellowship. R.A.F. is an Investigator of the Howard Hughes Medical Institute. The generation of mice for this work was supported by the Transgenic Core of the Yale DERC DK45735 and some of the work supported by a Pilot project from DK45735. This work was also supported by the JDRF.

**Author Contributions** E.E., S.H. and R.A.F. designed the study and wrote the manuscript; N.G. did the *in vitro* suppression assays and the flow analysis for IL-10 expression; A.E.H. and A.M.H. did the two-photon laser-scanning microscopy experiments; T.T. did the immunohistochemistry analysis; W.O. supported the work with key suggestions and by editing the manuscript; E.E. and S.H. did all other *in vitro* and *in vivo* experimental work; Y.Y.W. provided Foxp3-mRFP mice; A.R. did the viral infection experiments; N.V.R. provided clodronate-loaded liposomes, V.K.K. provided 2D2 mice and feedback on the manuscript; Y.I. provided *Il17a*<sup>-/-</sup> mice and feedback on the manuscript; J.K.K. provided the *Il17a*<sup>-/-</sup> mice and feedback on the manuscript and J.A.B. provided CD3-specific antibodies and feedback on the manuscript; K.C.H. provided teplizumab and key suggestions. E.E. and R.A.F. co-directed the project.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to E.E. ([enric.esplugues@yale.edu](mailto:enric.esplugues@yale.edu)) or R.A.F. ([richard.flavell@yale.edu](mailto:richard.flavell@yale.edu)).

## METHODS

**Mice.** BALB/c mice (blastocyst donors), CD1 mice (foster mothers), Tet-Cre transgenic mice ("deletor" mice, C57BL/6 background), C57BL/6 mice (B6), C57BL/6.Ly5.1 mice (CD45.1<sup>+</sup>), *IL6*<sup>-/-</sup> mice and *CCR6*<sup>-/-</sup> mice were purchased from The Jackson Laboratories. MOG-transgenic mice (2D2 mice, C57BL/6 background)<sup>30</sup> and Foxp3 reporter mice (FIR mice, C57BL/6 background)<sup>31</sup> were intercrossed with the IL-17A-eGFP reporter mice. We also used *Il17a*<sup>-/-</sup>, *Il17ra*<sup>-/-</sup> and IL-10-eGFP mice (Tiger mice)<sup>32-34</sup>. All mice were kept under specific pathogen-free conditions in the animal care facility at Yale University. The mice were studied at 6–12 week of age. All the experiments were approved by the Institutional Animal Care and Use Committee of Yale University.

**Generation of IL-17A-IRES-eGFP reporter mice.** A BAC clone consisting of *Il17a* genomic DNA derived from C57BL/6 mice was purchased from BacPac (Oakland, CA). An 8-kb BamHI-MluI fragment comprising exons 1, 2 and 3 for the *Il17a* gene was cloned into pEasy-Flox vector adjacent to the thymidine kinase selection marker. The internal ribosome entry site (IRES)-eGFP cassette was linked to a LoxP-flanked neomycin (Neo) selection marker to obtain the IRES-eGFP-Neo cassette. The targeting construct was generated by cloning the IRES-eGFP-Neo cassette into a SacII site between the translation stop codon (UGA) and the polyadenylation signal (A2UA3) of the *Il17a* gene. The targeting construct was linearized by ClaI cleavage and subsequently electroporated into Bruce4 C57BL/6 embryonic stem (ES) cells. Transfected ES cells were selected in the presence of 300 µg ml<sup>-1</sup> G418 and 1 µM ganciclovir. Drug-resistant ES cell clones were screened for homologous recombination by PCR. To obtain chimaeric mice, correctly targeted ES clones were injected into BALB/c blastocysts, which were then implanted into CD1 pseudopregnant foster mothers. Male chimaeras were bred with C57BL/6 to screen for germ-line transmitted offspring. Germ-line transmitted mice were bred with germline Cre transgenic mice (Tet-Cre mice) to remove the neomycin gene. Mice bearing the targeted *Il17a* allele were screened by PCR (*Il17a* knock-in (KI) sense: 5'-CACCAGCGCTGTGTCAT-3', *Il17a* KI anti-sense: 5'-ACAAACACGAAGCAGTTTGG-3' and *Il17a* IRES: 5'-ACCGGCCTTATTCGAAGC-3').

**Antibodies, tetramers and intracellular cytokine staining.** Anti-CD4 (L3T4), anti-CD62L (MEL-14), anti-CD44 (IM7), anti-CD45.1 (A20) and anti-CD45.2 (104), anti-TCRβ, anti-IL-2, anti-IL-17A, anti-TNFα, anti-Ki-67 and anti-BrdU (5'-bromo-2-deoxyuridine) were purchased from Becton Dickinson Pharmingen. For intracellular cytokine staining, the cells were restimulated with phorbol 12-myristate 13-acetate (PMA) (Sigma, 20 ng ml<sup>-1</sup>) and ionomycin (Sigma, 0.5 µg ml<sup>-1</sup>) for 4 h. GolgiStop (BD Bioscience) was added during the last 3 h of restimulation. After restimulation, the cells were washed and a Ficoll gradient was performed. The cells were fixed with 1% paraformaldehyde (electron microscopy grade) for 10 min on ice. After two washes, cells were incubated with FITC-conjugated anti-GFP antibody (Rockland) and phycoerythrin (PE)-conjugated anti-IL-17A (BD Bioscience) in wash/permeabilization solution (BD Bioscience) for 30 min on ice. Cells were washed twice and resuspended in PBS. Acquisitions were made with a LSRII cytometer (BD Bioscience).

For *ex vivo*-staining with MOG<sub>35-49</sub>/I-A(b)-tetramer-allophycocyanin (APC)-labelled (mouse myelin oligodendrocyte glycoprotein 38-49, "GWYRSPFSRWH", NIH Tetramer Facility), single-cell suspensions were incubated at a density of 10<sup>7</sup> cells ml<sup>-1</sup> with neuraminidase (0.7 µU ml<sup>-1</sup>, neuraminidase type X from *Clostridium perfringens*, Sigma) in serum-free DMEM at 37 °C/10% CO<sub>2</sub> for 25 min before incubation with the I-A(b) multimers (30 µg ml<sup>-1</sup>) in DMEM supplemented with 2% FCS (pH 8.0) at room temperature for 4 h. After washing, cells were stained for 7-AAD (Molecular Probes), CD4 (RM4-5) and TCRβ. hCLIP/I-A(b)-tetramer-APC-labelled was used as a control ("PVSKMRMATPLLMQA", NIH Tetramer Facility). The percentage of tetramer cells was determined in the CD4/TCRβ gate of live (7-AAD<sup>-</sup>) cells. Stained cells were analysed on LSRII cytometer (BD Bioscience) and data were analysed with FlowJo software (Treestar).

**Flow cytometry and FACS sorting.** Collected lymphocytes were treated with ammonium chloride lysis buffer (BioSource International) to remove red blood cells and washed with RPMI containing 10% FBS (Gemini Biological Products). Cells were then stained with a 1:400 dilution of the indicated antibodies together with 10 µg ml<sup>-1</sup> anti-Fc-Receptor blocking antibody (2.4G2, American Type Culture Collection) in PBS containing 2% FBS and then washed twice with PBS. For isolating T cells, CD4<sup>+</sup> T cells were first enriched by magnetic-activated cell sorting beads (Miltenyi Biotec) and then stained with the indicated antibodies. The Becton Dickinson FACS Vantage system and MoFlo sorter (DAKO Cytomation) were used for fluorescence detection and cell sorting.

**TH17 differentiation in vitro.** Splenocytes from IL-17A-IRES-eGFP mice and C57BL/6 mice were incubated with CD4-microbeads and then positively selected through LS columns (Miltenyi Biotec). After enrichment, naive cells (CD4<sup>+</sup> CD25<sup>-</sup> CD62L<sup>hi</sup> CD44<sup>low</sup>) were sorted by FACS as mentioned above. CD4<sup>+</sup> naive T cells were grown for 5 days at 10<sup>6</sup> cells ml<sup>-1</sup> with plate-bound anti-CD3 (5 µg

ml<sup>-1</sup>) and soluble anti-CD28 (2 µg ml<sup>-1</sup>) in medium (Bruff's medium supplemented with 10% FCS, L-glutamine, penicillin and streptomycin) under TH17 conditions (TGF-β, IL-6, IL-23, anti-IFN-γ, anti-IL4). IL-17A (eGFP) expression was determined by flow cytometry.

**Multiphoton imaging.** The small intestine (duodenum) from an IL-17A-eGFP × Foxp3-mRFP double reporter mouse treated with CD3-specific antibody was mounted on a glass slide in a chamber consisting of a silicone isolator (20 mm diameter × 0.5 mm, Electron Microscopy Sciences). The tissue was immersed in PBS and covered by a glass coverslip. An Olympus BX50WI microscope equipped with a ×20x, numerical aperture 0.95 Olympus objective and a LaVision TriMScope Multiphoton System controlled by Inspector Software (LaVision Biotec) was used to collect images. For excitation, a Coherent Chameleon Ti:Sapphire laser was tuned to 960 nm. Images of 300 × 300 µm size were recorded at a resolution of 1,024 × 1,024 pixels with 1-µm z-spacing. Emitted light was collected with nondescanned detectors after having passed 435/90, 525/50 and 615/100 nm bandpass filters. Velocity software (Improvizion) was used to create three-dimensional image stacks, and QuickTime Pro was used to generate image sequences.

**In vivo T-cell stimulation and intestinal lymphocyte isolation.** Different mice in C57BL/6 background were injected with anti-CD3 (20 µg, 145 2C11)<sup>34,35</sup> intraperitoneally 1–3 times at an interval of 2 days between injections and killed 4 h after the final injection. For the controls, isotype control or PBS was injected. The intraepithelial lymphocytes (IEL) and lamina propria lymphocytes (LPL) were collected as described with some modifications<sup>34</sup>. In brief, small or large intestines were removed and Peyer's patches were dissected. The first 2 cm of the small intestine were considered as duodenum. Intestines were opened longitudinally and then were cut into strips 1 cm in length. Tissues were washed with Hank's buffered saline and incubated in the presence of 5 mM of EDTA at 37 °C for 30 min. The released cells were loaded onto a Percoll gradient and centrifuged. The cells between 40% and 100% Percoll were collected and used as intestinal epithelial lymphocytes. LPL were collected by digesting gut tissue, which was removed for IEL isolation as described above. The tissue was digested with collagenase IV (100 U, Sigma) at 37 °C for 1 h and loaded onto a Percoll gradient and centrifuged. The cells between 40% and 100% Percoll were collected and used as LPL.

For the lumen content isolation and analysis, mice were anesthetized using isoflurane. One ligation was made after the pylorus and a second one about 4–5 cm distal from the first ligation. A small incision without breaching the vessel proximal to the second ligation was made and 10 ml of pre-warmed PBS (2 ml min<sup>-1</sup>) was injected using a syringe and a 27G1/2 needle. The fluids were collected in a Petri dish placed under the incision proximal to the second ligation. The collected fluids were incubated for 15 min in HBSS/ EDTA and then filtered through a 70 µm cell strainer before FACS analysis.

**Adoptive transfer of CD4<sup>+</sup> T cells.** CD4<sup>+</sup>FoxP3<sup>+</sup> and the CD4<sup>+</sup>FoxP3<sup>-</sup> T cells from the thymus or from the periphery of IL17A-eGFP × Foxp3-mRFP double reporter mice were collected and purified by magnetic-activated cell sorting (MACS; Miltenyi Biotec). After MACS enrichment, total CD4<sup>+</sup>FoxP3<sup>+</sup> and CD4<sup>+</sup>FoxP3<sup>-</sup> T cells were FACS-sorted and 4 × 10<sup>6</sup> T cells were adoptively transferred (intravenously) into sub-lethally irradiated, sex-matched wild-type CD45.1<sup>+</sup> recipient mice. Four weeks after transfer, animals were injected with CD3-specific antibody (20 µg) and the small intestines were recovered and examined for eGFP and mRFP by FACS.

**Immunofluorescence microscopy.** Small intestines were removed from IL-17A-eGFP reporter mice and wild-type littermates after CD3-specific antibody treatment *in vivo*. Small intestines were fixed in 4% paraformaldehyde for 16 h. After two washes with PBS 20% sucrose solution was added. The 20% sucrose solution was replaced 16 h later with 30% sucrose solution. On the next day, the samples were washed and then snap-frozen in OCT and stored at -80 °C. Cryosections were cut at 12 µm on a Leica model CM1850 freezing microtome, applied to Superfrost Plus Gold slides (Fisher Scientific), air-dried, and PAP pen applied (Zymed Laboratories). Sections were blocked for 30 min at ambient temperature with serum-free protein block (Dako Cytomation) and were stained with PE-anti-CD4 (BD) and Alexa488-anti-GFP (Invitrogen) overnight at 4 °C. Samples were washed three times by immersing in PBS for 5 min and then mounted with Prolong gold mounting media with DAPI (Invitrogen). Sections were observed under dark field in independent fluorescence channels using an automated Olympus BX-61 microscope.

**Experimental autoimmune encephalomyelitis.** Mice were immunized subcutaneously with 250 µg of MOG<sub>35-55</sub> (Yale Keck facility) emulsified in CFA (BD Difco). Mice received 400 ng pertussis toxin (PTx, List Biological Laboratories) intraperitoneally at the time of immunization and 48 h later. Mice were checked for clinical symptoms daily, and signs were translated into clinical score as follows: 0, no detectable signs of EAE; 0.5, tail weakness; 1, complete tail paralysis; 2, partial hind limb paralysis; 2.5, unilateral complete hind limb paralysis; 3, complete



bilateral hind limb paralysis; 3.5, complete hind limb paralysis and partial forelimb paralysis; 4, total paralysis of forelimbs and hind limbs (mice with a score above 3.5 to be killed); 5, death. All animal experiments were conducted according to the IACUC policies.

**Cytokine assays.** Cytokines were quantified in plasma by ELISA (TGF- $\beta$ 1, Promega) or by Cytometric Bead Array (IL-6 and IL-17A, BD Bioscience) following the manufacturer's instructions. The plasma was obtained by centrifugation of blood collected on EDTA-coated tubes after cardiac puncture.

**Gene expression analysis.** Total RNA extracted (100 ng; RNeasy, Qiagen) from intestinal rT<sub>H</sub>17 cells (from CD3-specific antibody-treated animals) or from pro-inflammatory T<sub>H</sub>17 cells (from EAE-induced mice) were used to perform a genome-wide transcriptional profiling assay (GeneChip Mouse 1.0 ST Array, Affymetrix). Data was analysed with GeneSpring GX 10 (Agilent Technologies).

**Relative gene expression analysis.** RNA from cells/tissues was isolated with the RNeasy/QIAshredder purification system (Qiagen) in accordance with the manufacturer's protocol. RNA was subjected to reverse transcriptase with Superscript II (Invitrogen) with oligo(dT) primer in accordance with the manufacturer's protocol. cDNA was semi-quantified using commercially available primer/probe sets (Applied Biosystems) and analysed with the  $\Delta\Delta C_t$  (change in cycle threshold) method. All results were normalized to *Hprt* quantified in parallel amplification reactions during each PCR quantification.

**Suppression assays.** CFSE (2  $\mu$ M)-labelled CD4<sup>+</sup>CD25<sup>-</sup> T cells (responder cells) were cultured in 96-well round bottom plates at  $2 \times 10^4$  cells per well with  $10^5$  irradiated APCs (spleenocytes MACS-depleted for CD4<sup>+</sup> and CD8<sup>+</sup> T cells) as feeder cells in the presence of  $2 \times 10^4$  cells per well of FACS-sorted CD4<sup>+</sup>IL-17A<sup>+</sup>Foxp3<sup>-</sup> or CD4<sup>+</sup>IL-17A<sup>-</sup>Foxp3<sup>+</sup> T cells. Cell cultures were stimulated with 2  $\mu$ g ml<sup>-1</sup> of anti-CD3 antibody (2C11) in the presence or not of anti-TGF- $\beta$  (1D11), anti-CTLA-4 (9H10) and anti-IL10R. After 4 days, cells were collected, stained and the CFSE signal was analysed by flow cytometry.

**Sepsis induced by infection and superantigen treatment:** *Staphylococcus aureus* (ATCC 14458, SEB<sup>+</sup> TSST-1<sup>-</sup>) was injected intravenously into IL-17A-eGFP reporter mice ( $10^8$  colony-forming units per mouse). Mice were killed 3 days after the injection, at a time when they displayed severe clinical symptoms of sepsis (weight loss, dehydration, lethargy) and the presence of CD4<sup>+</sup>IL17A<sup>+</sup> T cells was tested in different organs (spleen, lymph node, small intestine) using FACS analysis. Similar experiments were done injecting the superantigens SEB and TSST-1. All of them were purchased from Toxin Technology. All superantigens were administered three times (0, 48, 96 h) intraperitoneally at 50  $\mu$ g per mouse.

For the influenza A infection, mice were infected with  $1 \times 10^4$  plaque-forming units of influenza A/PR8 (H1N1) virus via the intranasal route. Infection was

performed by the intranasal application of 50  $\mu$ l virus stock diluted in PBS (or an equal volume of PBS as a control) to mice that had been deeply anesthetized with anafane (Ivesco). Lungs and small intestines were harvested 3 and 5 days after infection for flow cytometry analysis.

**Peripheral blood mononuclear cells isolation and administration.** Human leukocytes were collected by leukapheresis of adult volunteer donors under a protocol approved by the Yale Human Investigations Committee. The peripheral blood mononuclear cells were isolated using Lymphocyte Separation Medium (Cappel) according to the manufacturer's instructions. The cells were stored in 10% DMSO/90% FBS at  $-196^\circ\text{C}$  and were thawed and washed before use. *Rag2*<sup>-/-</sup>  $\times$   *$\gamma$ c*<sup>-/-</sup> double knockout mice were reconstituted with  $5 \times 10^7$  human peripheral blood mononuclear cells by intraperitoneally inoculation 2 weeks before anti-CD3 specific antibody treatment. The number of human T cells (CD45<sup>+</sup>CD4<sup>+</sup>) in the small intestine was evaluated by flow cytometry. Animals demonstrated no signs of graft-vs-host disease. Rare animals that failed to reconstitute with human T cells were, by prior design, excluded from analysis.

**Endoscopic procedure.** Colonoscopy was performed in a blinded fashion for colitis scoring using the Coloview system (Karl Storz, Germany). Briefly: colitis scoring was based on granularity of mucosal surface, stool consistence, vascular pattern, translucency of the colon and fibrin visible (0–3 points for each).

**Statistical analysis.** Where indicated, the Student *t* test for non-paired data and the Mann–Whitney *U* test were used to calculate statistical significance for differences in a particular measurement between different groups. A *P*-value of less than 0.05 was considered significant.

30. Bettelli, E. *et al.* Myelin oligodendrocyte glycoprotein-specific T cell receptor transgenic mice develop spontaneous autoimmune optic neuritis. *J. Exp. Med.* **197**, 1073–1081 (2003).
31. Wan, Y. Y. & Flavell, R. A. Identifying Foxp3-expressing suppressor T cells with a bicistronic reporter. *Proc. Natl Acad. Sci. USA* **102**, 5126–5131 (2005).
32. Nakae, S. *et al.* Antigen-specific T cell sensitization is impaired in IL-17-deficient mice, causing suppression of allergic cellular and humoral responses. *Immunity* **17**, 375–387 (2002).
33. Ye, P. *et al.* Requirement of interleukin 17 receptor signaling for lung CXC chemokine and granulocyte colony-stimulating factor expression, neutrophil recruitment, and host defense. *J. Exp. Med.* **194**, 519–528 (2001).
34. Kamanaka, M. *et al.* Expression of interleukin-10 in intestinal lymphocytes detected by an interleukin-10 reporter knockin *tiger* mouse. *Immunity* **25**, 941–952 (2006).
35. Alegre, M. L. *et al.* An anti-murine CD3 monoclonal antibody with a low affinity for Fc gamma receptors suppresses transplantation responses while minimizing acute toxicity and immunogenicity. *J. Immunol.* **155**, 1544–1555 (1995).

# A Polycomb-based switch underlying quantitative epigenetic memory

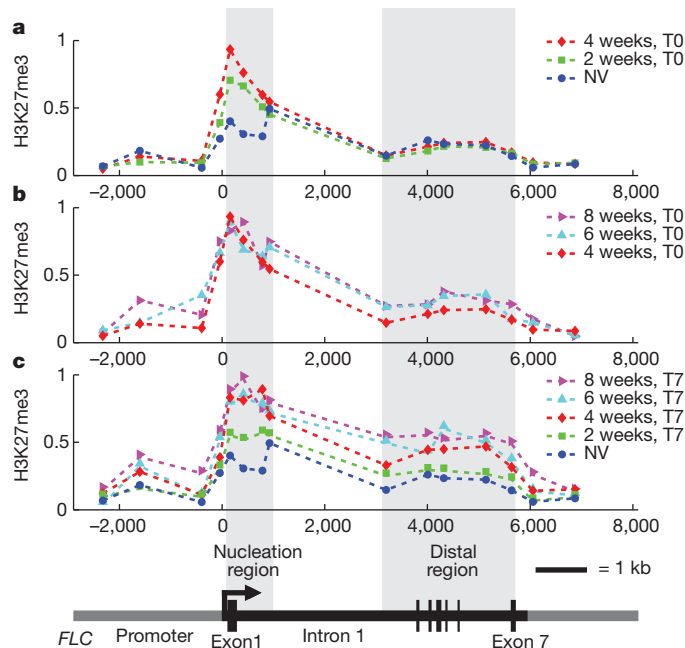
Andrew Angel<sup>1\*</sup>, Jie Song<sup>2\*</sup>, Caroline Dean<sup>2</sup> & Martin Howard<sup>1</sup>

The conserved Polycomb repressive complex 2 (PRC2) generates trimethylation of histone 3 lysine 27 (H3K27me3)<sup>1,2</sup>, a modification associated with stable epigenetic silencing<sup>3,4</sup>. Much is known about PRC2-induced silencing but key questions remain concerning its nucleation and stability. Vernalization, the perception and memory of winter in plants, is a classic epigenetic process that, in *Arabidopsis*, involves PRC2-based silencing of the floral repressor *FLC*<sup>5,6</sup>. The slow dynamics of vernalization, taking place over weeks in the cold, generate a level of stable silencing of *FLC* in the subsequent warm that depends quantitatively on the length of the prior cold. These features make vernalization an ideal experimental system to investigate both the maintenance of epigenetic states and the switching between them. Here, using mathematical modelling, chromatin immunoprecipitation and an *FLC*:GUS reporter assay, we show that the quantitative nature of vernalization is generated by H3K27me3-mediated *FLC* silencing in the warm in a subpopulation of cells whose number depends on the length of the prior cold. During the cold, H3K27me3 levels progressively increase at a tightly localized nucleation region within *FLC*. At the end of the cold, numerical simulations predict that such a nucleation region is capable of switching the bistable epigenetic state of an individual locus, with the probability of overall *FLC* coverage by silencing H3K27me3 marks depending on the length of cold exposure. Thus, the model predicts a bistable pattern of *FLC* gene expression in individual cells, a prediction we verify using the *FLC*:GUS reporter system. Our proposed switching mechanism, involving the local nucleation of an opposing histone modification, is likely to be widely relevant in epigenetic reprogramming.

In *Arabidopsis*, vernalization requires the plant-homeodomain-PRC2 complex (PHD-PRC2), whose components include SWINGER (an E(Z) histone methyltransferase homologue), together with VIN3, VRN5 and VEL1 (PHD proteins required to generate high H3K27me3 levels)<sup>7–10</sup>. PRC2 complexes such as this are thought to initiate trimethylation of H3K27 and then spread this modification through a positive feedback loop where they both bind to H3K27me3 and trimethylate other nucleosomes<sup>7–10</sup>. In principle, this process could generate stable maintenance of histone modifications even when perturbed by DNA replication, where on average half of the modifications could be lost by the insertion of unmodified nucleosomes. In general, however, how to establish stable epigenetic silencing at a level that depends quantitatively on the level of a transient stimulus is not understood.

In an effort to answer this question, we combined chromatin immunoprecipitation (ChIP) experiments with mathematical modelling to dissect the vernalization process. This required high-resolution analysis of H3K27me3 during vernalization as a function of space across the *FLC* locus and time. Previously, we had demonstrated that PRC2 was associated with the whole *FLC* locus before cold, and that this correlated with an epigenetically stable low level of silencing<sup>9,11</sup>. After 8 weeks of cold, a PHD-PRC2 complex had formed at one particular site within the gene and only after plants were returned to the warm did the

PHD proteins associate more generally with PRC2 over the locus. This association resulted in high levels of H3K27me3, which epigenetically maintained the fully silent state. Therefore, we needed to explore the quantitative accumulation of silencing during the cold. Figure 1a (and Supplementary Fig. 1) shows the H3K27me3 profile across *FLC*, in non-vernalized plants and also after 2 and 4 weeks of cold (at 4 °C). A localized peak of H3K27me3 develops (close to exon 1, see Fig. 1), the level of which increases with increasing length of cold, in agreement with previous results<sup>12</sup>. For longer periods of cold (6 and 8 weeks), this nucleation seems to saturate, with some limited general increase in H3K27me3 levels across the rest of the locus (Fig. 1b and Supplementary Fig. 2). During the cold, the rising levels of H3K27me3 in the nucleation region are correlated with an increase in expression of the PHD protein VIN3 (refs 7, 8), whose concentration levels may therefore constitute one of the primary read-outs of the length of cold. However, VIN3 expression returns to low levels immediately (within one day) after return to the warm: the memory of the cold implicit in the VIN3 concentration levels must therefore be converted into a



**Figure 1 | H3K27me3 ChIP experiments.** **a**, H3K27me3 profile across *FLC* for non-vernalized plants (NV) and after 2 or 4 weeks of cold, with 0 days of warm (T0). **b**, H3K27me3 profile across *FLC* after 4, 6 or 8 weeks of cold, with 0 days of warm (T0). **c**, H3K27me3 profile across *FLC* for non-vernalized plants and 7 days post-cold (T7) after 2, 4, 6 or 8 weeks of cold (top) and schematic layout of *FLC* gene (bottom), with definitions of nucleation and distal regions. Panels (a–c) show results from one complete experiment. Similar results were obtained with a biological replicate (Supplementary Figs 1–3).

<sup>1</sup>Department of Computational and Systems Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK. <sup>2</sup>Department of Cell and Developmental Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK.

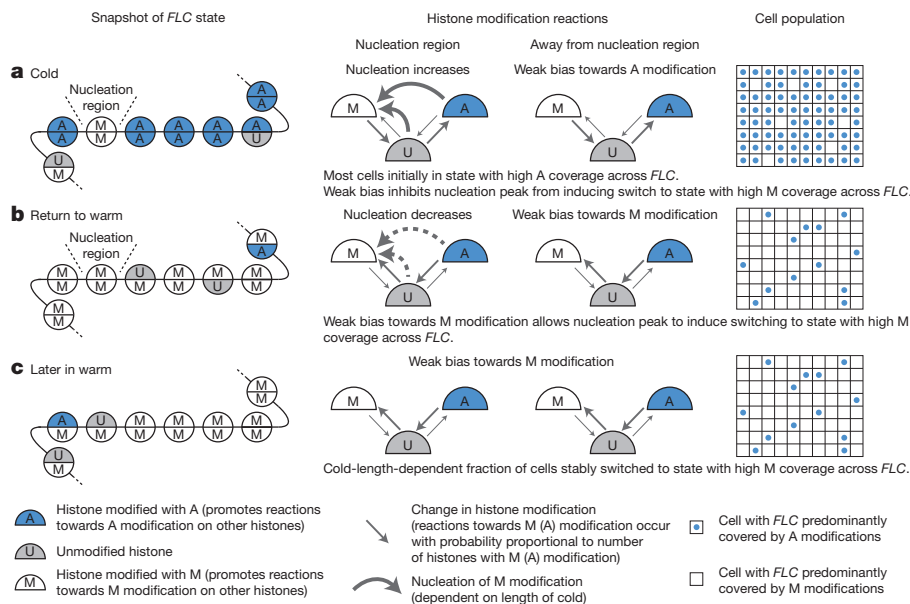
\*These authors contributed equally to this work.

second memory element in the warm, potentially H3K27me3 levels. We measured therefore the H3K27me3 profile across *FLC* 7 days after the plants had been transferred to the warm (Fig. 1c and Supplementary Fig. 3). The profile changed rather little in the nucleation region but rose quantitatively across the rest of the *FLC* locus according to the length of the cold period (Supplementary Fig. 4). This quantitative increase correlated with decreased *FLC* expression (Supplementary Fig. 5), and was consistent with H3K27me3 levels being a key element in epigenetic memory of the cold after return to the warm.

For shorter periods of cold, levels of ChIP-measured H3K27me3 at the nucleation region depended quantitatively on the length of the cold period, and could therefore potentially act to switch quantitatively the epigenetic state of the locus after return to the warm. Epigenetic states are, however, intrinsically stable (in the case of vernalization, over many weeks or months). It was therefore unclear whether the small size of the nucleation region (Fig. 1a) would be sufficient to cause a quantitative switch in the epigenetic state of the *FLC* locus. To answer this question, we turned to mathematical modelling<sup>13–19</sup>. Previous modelling of epigenetic states had focused primarily on cell-autonomous dynamics in yeast<sup>13,15–18</sup>. We wondered whether a similar cell-autonomous mechanism could also explain the quantitative response to the cold seen in vernalization. The above data provided an ideal opportunity to test such a model. In essence, our model is an implementation of a Polycomb-based switch, incorporating highly dynamic rewriting of histone modifications on a timescale of minutes (see Fig. 2 and Supplementary Information for details). The rapid dynamics of histone modifications are vital as they allow for the formation of bistable epigenetic states, despite the noisy turnover of nucleosomes, both continuously (on a timescale of tens of minutes<sup>20</sup>) and discretely at DNA replication (on a timescale of days<sup>21,22</sup>). The model utilizes a simplified set of possible histone modifications: M (H3K27me3), U (unmodified) and A (activating) (other possible assumptions are discussed in the Supplementary Information), where histones with the M or A modification antagonistically promote the transition of other histones towards the same modification, leading to positive feedback<sup>13,16</sup>. At the high levels of noise potentially generated by noisy nucleosome turnover and subsequent random histone mark addition, the presence of spatially long-ranged interactions between histones is an important

factor in the formation of bistable states<sup>13</sup>. In our model, starting from a state with the *FLC* locus covered by A modifications, two additional processes contribute to the switch-like behaviour at the end of the cold into the opposing state, where the locus is predominantly covered by M modifications. The first process is a site-specific nucleation of the opposing M modification mediated by a cold-induced increase of VIN3 expression (motivated by the localized increase of H3K27me3 levels, see Fig. 1a). The second process is a permanent PHD-PRC2-mediated bias in the histone dynamics towards the M modification on return to the warm, motivated by the observed post-cold spread of the PHD protein VRN5 across the locus<sup>9</sup>. The combination of these two processes enhances the probability of switching between the two bistable epigenetic states, before nucleation is lost (due to clearance of VIN3 on a timescale of days at the end of the cold) and the potential to switch is greatly diminished.

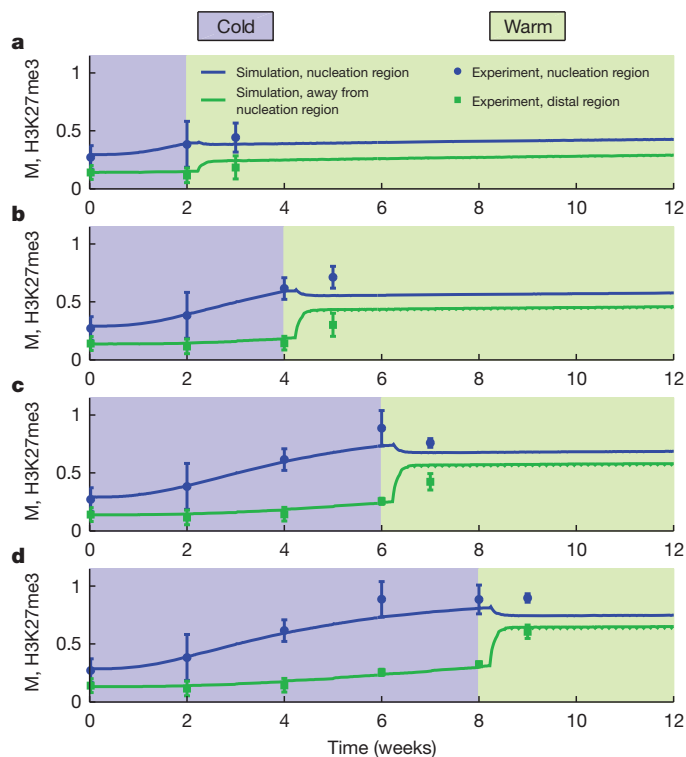
We simulated the above model in 10,000 individual *FLC* loci, and attempted to fit unconstrained model parameters to our experimental ChIP data (see Supplementary Table 2). In Fig. 3a–d we show the simulated population-averaged levels of the M modification averaged separately over the nucleation region (five-histones wide) and the rest of the locus. Also shown are the experimental H3K27me3 values for non-vernalized plants and 0 and 7 days after 2, 4, 6 or 8 weeks of cold, averaged separately over the nucleation and distal regions of *FLC* (as defined in Fig. 1). We find good overall agreement with the observed data. During the cold, some limited increase of H3K27me3 levels does occur in distal regions (Fig. 1b). However, by tracking the dynamics of individual simulated *FLC* loci, we predict that these low levels of H3K27me3 arise from a large subpopulation of cells with little or no H3K27me3 coverage away from the nucleation region and a small subpopulation where H3K27me3 substantially covers the locus (occupying ~90% of histones, limited by noisy dynamics such as nucleosome turnover). Switching into this latter subpopulation is limited by the lack of a (VRN5-dependent) bias towards the M modification. On return to the warm, however, this limitation is removed: our simulations then demonstrate that a transient narrow peak of H3K27me3 can indeed cause the epigenetic state of the locus to be efficiently switched. Again, this rise in H3K27me3 levels is predicted to be due to a subpopulation of cells with substantial coverage of H3K27me3 across the



**Figure 2 | Schematic outline of mathematical model for *FLC* silencing.** a–c, Left, snapshot of *FLC* histone state; middle, histone modification reactions; right, cell population status. a, During the cold, M modifications are added at the nucleation region; most cells are still locked in a state with high A modification coverage. b, At the end of the cold, nucleation dies away, but

histone modification reactions are biased towards M modification. This leads to a short window in which the system can switch from high A to high M modification coverage. c, After the cold, a subpopulation of cells, whose number depends quantitatively on the length of the cold, stably switched to a silenced state, with high levels of M modification coverage.





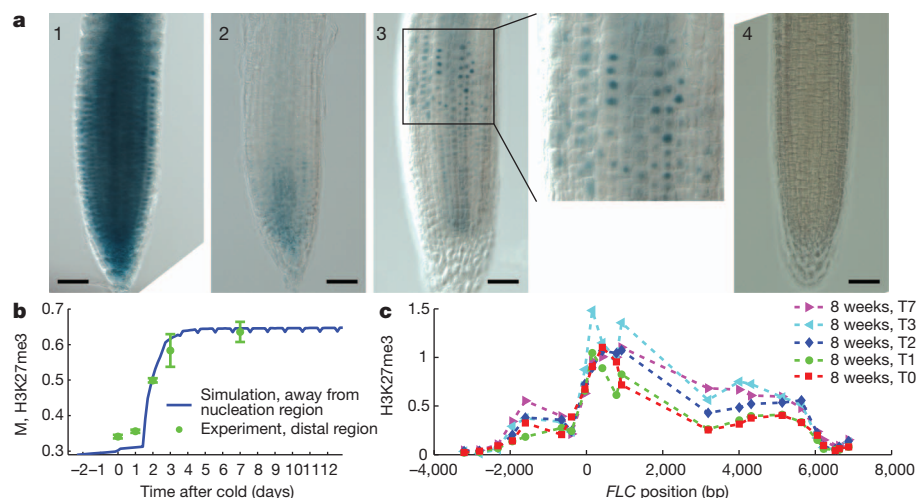
**Figure 3 | Fitting model output to experimental ChIP data.** **a–d**, Simulated population-averaged levels of M modification, averaged separately at and away from the nucleation region, as compared to experimental values of H3K27me3, averaged separately over nucleation and distal regions, for non-vernalized plants and at 0 and 7 days post-cold after 2 (**a**), 4 (**b**), 6 (**c**) or 8 (**d**) weeks of cold. Error bars represent the range,  $n = 2$ .

*FLC* locus, with the remaining cells having low H3K27me3 levels. Simulated population-averaged levels of the M modification from our fitted simulations are found to be approximately stable up to 30 days after the cold (Fig. 3), with a population-level change of less than about 10% from 7 days after the cold, in accordance with the stable long-term silencing observed experimentally. Simulated *FLC* expression levels 7 days after the cold also show a robust reduction from

non-vernalized levels, a reduction whose magnitude increases with increasing length of cold (Supplementary Fig. 5). However, with our fitted parameters, the simulated reduction is not as large as seen experimentally. The model therefore indicates that additional factors could reinforce repression of *FLC*, consistent with a role for the DNA-binding protein VRN1 (ref. 23).

According to the model, quantitative silencing on return to the warm is achieved through the fraction of cells that switch to the silenced state, a fraction that increases with increasing length of cold. Although the active or silenced epigenetic states are stable, a nucleating peak gives rise to a small probability per unit time of switching. This switching is generated by the noisy histone modification dynamics, which are able to amplify the signal from the nucleating peak of H3K27me3 and use a rare fluctuation to flip the epigenetic state of the system. The timescale of a switch at an individual locus from predominant A to M coverage is predicted to be very rapid (about 45 min, see Supplementary Fig. 6).

To test the model prediction of bistable *FLC* expression, we assayed expression of an *FLC:GUS* translational fusion<sup>24,25</sup> in tissue exposed to different cold treatments. For ease of imaging we analysed root tissue, in which cold-induced epigenetic silencing of *FLC* occurs, and which can regenerate into plants that have ‘remembered’ the cold exposure<sup>26</sup>. The results are shown in Fig. 4a and Supplementary Fig. 8, and show that without cold the gene is evenly expressed in the dividing cells of the root tissue. This expression is reduced in all cells after 2 weeks of cold, supporting the fast reduction in *FLC* transcription seen previously<sup>27</sup>. However, on return to the warm, the expression reactivates in a proportion of cells, in a cell-autonomous manner. These cells reflect the proportion of the population where the epigenetic switch to the silent state has not occurred. This heterogeneity in silencing would explain the flowering variation in regenerants from partially vernalized plants<sup>28</sup>. The proportion of cells reactivating *FLC* expression reduces with increasing time in the prior cold (Fig. 4a), fully supporting the model predictions. The diversity of sector patterns reflects the stochastic nature of the silencing process: both single cells and clusters of *FLC:GUS*-expressing cells are observed, the latter probably reflecting clones from a single reactivating cell (Supplementary Fig. 8). The modelling also predicted that population-averaged levels of H3K27me3 could rise rapidly (within 1–2 days) after return to the warm (Fig. 4b). Such increasing levels reflect individual *FLC* loci undergoing very rapid



**Figure 4 | Validating model predictions.** **a**, *Arabidopsis* roots expressing *FLC:GUS* from plants with various cold treatments. Panels from left to right: (1) non-vernalized plants; (2) cold-treated for 2 weeks, harvested immediately on return to warm (T0); (3) cold-treated for 2 weeks, then grown for a further 7 days in the warm (T7); or (4) cold-treated for 4 weeks and then grown for a further 5 days in the warm (T5). Scale bars, 50  $\mu$ m. **b**, Simulated population-averaged levels of M modification, averaged away from nucleation region, as

compared to experimental values of H3K27me3 (from **c** and Supplementary Fig. 9), averaged over distal region 0, 1, 2, 3 or 7 days post-cold after 8 weeks of cold. Error bars represent the range,  $n = 2$ . **c**, Higher time resolution ChIP data for H3K27me3 profile across *FLC* at 0, 1, 2, 3 or 7 days post-cold after 8 weeks of cold. Similar results were obtained with a biological replicate (Supplementary Fig. 9).

switching (taking only about 45 min) into the silenced epigenetic state, but at different starting times. To experimentally test the prediction of a rapid rise at population level, we performed additional higher time resolution (daily) ChIP experiments for the week immediately following the cold (Fig. 4c and Supplementary Fig. 9). In accordance with model predictions, we found a rapid rise of population-averaged H3K27me3 levels, with the H3K27me3 profile increasing mostly between 1 and 3 days after the cold.

Our combined experimental/modelling approach to vernalization has revealed that the quantitative level of epigenetic silencing in the warm is generated by an appropriate proportion of cells switching into a stable *FLC*-silenced state. Controlled switching of individual cells between bistable states, mediated by nucleation of an opposing histone mark may therefore be a common method of achieving a quantitative epigenetic response at a population level.

## METHODS SUMMARY

The ChIP experiments and *FLC* expression analysis were carried out using *Col Fri-sf2* (ref. 29). H3K27me3 and H3 levels were assayed using anti-H3K27me3 (Millipore 07-449) and anti-H3 (Abcam 1791), respectively. *SHOOT MERISTEMLESS (STM)*<sup>30</sup> was used as the internal control for the ChIP experiments (primers are listed in Supplementary Table 1). Data are represented as the ratio of (H3K27me3 *FLC*/H3 *FLC*) to (H3K27me3 *STM*/H3 *STM*). *FLC::GUS* assay was performed using *FLC::Cok::GUS* in *Ler FRI*<sup>25</sup>. Photoshop adjustment involved only image exposure using adjustment levels.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 27 January; accepted 26 May 2011.

Published online 24 July 2011.

- Hansen, K. H. *et al.* A model for transmission of the H3K27me3 epigenetic mark. *Nature Cell Biol.* **10**, 1291–1300 (2008).
- Margueron, R. *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**, 762–767 (2009).
- Turner, B. M. Defining an epigenetic code. *Nature Cell Biol.* **9**, 2–6 (2007).
- Margueron, R. & Reinberg, D. Chromatin structure and the inheritance of epigenetic information. *Nature Rev. Genet.* **11**, 285–296 (2010).
- Sheldon, C. C. *et al.* The *FLF* MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. *Plant Cell* **11**, 445–458 (1999).
- Michaels, S. D. & Amasino, R. M. *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* **11**, 949–956 (1999).
- Sung, S. & Amasino, R. M. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature* **427**, 159–164 (2004).
- Greb, T. *et al.* The PHD finger protein VRN5 functions in the epigenetic silencing of *Arabidopsis FLC*. *Curr. Biol.* **17**, 73–78 (2007).
- De Lucia, F., Crevillen, P., Jones, A. M., Greb, T. & Dean, C. A. PHD-Polycomb repressive complex 2 triggers the epigenetic silencing of *FLC* during vernalization. *Proc. Natl Acad. Sci. USA* **105**, 16831–16836 (2008).
- Wood, C. C. *et al.* The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. *Proc. Natl Acad. Sci. USA* **103**, 14631–14636 (2006).
- Gendall, A. R., Levy, Y. Y., Wilson, A. & Dean, C. The *VERNALIZATION 2* gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. *Cell* **107**, 525–535 (2001).
- Finnegan, E. J. & Dennis, E. S. Vernalization-induced trimethylation of histone H3 lysine 27 at *FLC* is not maintained in mitotically quiescent cells. *Curr. Biol.* **17**, 1978–1983 (2007).
- Dodd, I. B., Micheelsen, M. A., Sneppen, K. & Thon, G. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell* **129**, 813–822 (2007).
- Salazar, J. D., Foreman, J., Carr, I. A., Rand, D. A. & Millar, A. J. Mathematical model of the epigenetic control of vernalisation in *Arabidopsis thaliana*. *Acta Hort. (ISHS)* **803**, 187–192 (2008).
- Sedighi, M. & Sengupta, A. M. Epigenetic chromatin silencing: bistability and front propagation. *Phys. Biol.* **4**, 246–255 (2007).
- David-Rus, D., Mukhopadhyay, S., Lebowitz, J. L. & Sengupta, A. M. Inheritance of epigenetic chromatin silencing. *J. Theor. Biol.* **258**, 112–120 (2009).
- Mukhopadhyay, S., Nagaraj, V. H. & Sengupta, A. M. Locus dependence in epigenetic chromatin silencing. *Biosystems* **102**, 49–54 (2010).
- Kelemen, J. Z., Ratna, P., Scherrer, S. & Becskei, A. Spatial epigenetic control of mono- and bistable gene expression. *PLoS Biol.* **8**, e1000332 (2010).
- Kaufman, P. D. & Rando, O. J. Chromatin as a potential carrier of heritable information. *Curr. Opin. Cell Biol.* **22**, 284–290 (2010).
- Deal, R. B., Henikoff, J. G. & Henikoff, S. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* **328**, 1161–1164 (2010).
- Reddy, G. V., Heisler, M. G., Ehrhardt, D. W. & Meyerowitz, E. M. Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of *Arabidopsis thaliana*. *Development* **131**, 4225–4237 (2004).
- Grandjean, O. *et al.* In vivo analysis of cell division, cell growth, and differentiation at the shoot apical meristem in *Arabidopsis*. *Plant Cell* **16**, 74–87 (2004).
- Levy, Y. Y., Mesnage, S., Mylne, J. S., Gendall, A. R. & Dean, C. Multiple roles of *Arabidopsis VRN1* in vernalization and flowering time control. *Science* **297**, 243–246 (2002).
- Bastow, R. *et al.* Vernalization requires epigenetic silencing of *FLC* by histone methylation. *Nature* **427**, 164–167 (2004).
- Sheldon, C. C. *et al.* Resetting of *FLOWERING LOCUS C* expression after epigenetic repression by vernalization. *Proc. Natl Acad. Sci. USA* **105**, 2214–2219 (2008).
- Burn, J. E., Bagnall, D. J., Metzger, J. D., Dennis, E. S. & Peacock, W. J. DNA methylation, vernalization, and the initiation of flowering. *Proc. Natl Acad. Sci. USA* **90**, 287–291 (1993).
- Swiezewski, S., Liu, F., Magusin, A. & Dean, C. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature* **462**, 799–802 (2009).
- Pierik, R. L. M. in *Cellular and Molecular Aspects of Floral Induction* (ed. Bernier, G.) 409–415 (Longmans, 1970).
- Lee, I., Michaels, S. D., Masshardt, A. S. & Amasino, R. M. The late-flowering phenotype of *FRIGIDA* and mutations in *LUMINIDEPENDENS* is suppressed in the Landsberg erecta strain of *Arabidopsis*. *Plant J.* **6**, 903–909 (1994).
- Zhang, X. *et al.* Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.* **5**, e129 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank all members of the C.D. and M.H. groups for discussions. We also thank S. Costa for suggestions to improve the *FLC::GUS* imagery and V. Grieneisen, S. Maree, R. Morris, S. Swiezewski and P. Wigge for comments on the manuscript. This research was supported by an Advanced Investigator European Research Council grant and the Core Strategic Grant from the Biotechnology and Biological Sciences Research Council to the John Innes Centre. M.H. also acknowledges support from The Royal Society.

**Author Contributions** C.D. and M.H. conceived the study, A.A., J.S., C.D. and M.H. designed the experiments, J.S. performed the experiments, A.A. and J.S. analysed the experimental data, A.A. and M.H. designed the numerical simulations, A.A. performed the simulations and analysed the simulation data. A.A., J.S., C.D. and M.H. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to C.D. ([caroline.dean@jic.ac.uk](mailto:caroline.dean@jic.ac.uk)) or M.H. ([martin.howard@jic.ac.uk](mailto:martin.howard@jic.ac.uk)).

## METHODS

**Plant material and growth conditions.** The *Columbia* line *FRI*-Sf2 was described previously<sup>29</sup>. Plant growth conditions were described previously<sup>9</sup>. Plants were vernalized for 2, 4, 6 or 8 weeks for the ChIP experiments, together with the non-vernalized control. T0 seedlings were harvested immediately after prolonged cold, whereas T1, T2, T3 and T7 seedlings were grown for 1, 2, 3 and 7 days, respectively, after transfer back to the warm.

**ChIP and real-time quantitative PCR analysis.** ChIP assays were performed as previously described<sup>9</sup> with modifications. Trimethyl-histone H3 lysine 27 was assayed using anti-trimethyl-histone H3 lysine 27 from Millipore/Upstate (catalogue no. 07-449). Histone H3 levels were assayed using anti-H3 core antibody from Abcam (catalogue no. 1791). After immunoprecipitation, DNA was recovered using Chelex 100 resin (Bio-Rad, 10 g per 100 ml ddH<sub>2</sub>O). All ChIP experiments were quantified by quantitative PCR (qPCR) in triplicates with appropriate primers (Supplementary Table 1). *SHOOT MERISTEMLESS* (*STM*)<sup>30</sup> was used as the internal control for the ChIP experiments. Data are represented as the ratio of (H3K27me3 *FLC*/H3 *FLC*) to (H3K27me3 *STM*/H3 *STM*).

**Expression analysis.** cDNA was synthesized using SuperScript III (Invitrogen) with a mixture of oligo d(T) and gene-specific primer (*FLC\_unspliced\_R* 5'-cttgtaatcaaagtg gagagc-3'), and analysed by qPCR on a LightCycler 480 II instrument (Roche), using LightCycler 480 SYBR Green mix (Roche). *FLC* sense unspliced (unspliced intron 2/3, *FLC\_unspliced\_F* 5'-cgcaatttcatagcccttg-3', *FLC\_unspliced\_R*) quantification was normalized against the *Arabidopsis* *UBC* gene (*At5g25760*, *UBC\_F* 5'-ctgcgactcagg gaatcttctaa-3', *UBC\_R* 5'-ttgtgccattgaattgaacc-3').

***FLC:GUS* assay.** The *FLC-Col:GUS* in *Ler FRI* was described previously<sup>25</sup>. The *FLC-Col:GUS* translational fusion contains the  $\beta$ -glucuronidase coding sequences cloned into the *NheI* site within *Col FLC* exon 6 (ref. 24). GUS staining was carried out by 4.5 h incubation at 37 °C in X-Gluc staining solution (0.5 mg ml<sup>-1</sup> X-Gluc, 2.0 mM K<sub>3</sub>Fe(CN)<sub>6</sub>, 2.0 mM K<sub>4</sub>Fe(CN)<sub>6</sub>, 100 mM phosphate buffer pH 7.0, 0.1% [v:v] Triton X-100, 10 mM EDTA). Seedlings were vacuum infiltrated for 5 min before incubation. After staining, seedlings were cleared in 75% (v:v) ethanol and then 8:1:3 (w:w:w) chloral hydrate:glycerol:water. Roots were imaged using DIC optics on a Leica DM6000 microscope with a Leica DFC420 camera controlled via Leica LAS AF7000 software. Photoshop adjustment involved only image exposure using adjustment levels.



# Neuronal basis of age-related working memory decline

Min Wang<sup>1</sup>, Nao J. Gamo<sup>1</sup>, Yang Yang<sup>1</sup>, Lu E. Jin<sup>1</sup>, Xiao-Jing Wang<sup>1</sup>, Mark Laubach<sup>1,2</sup>, James A. Mazer<sup>1</sup>, Daeyeol Lee<sup>1</sup> & Amy F. T. Arnsten<sup>1</sup>

Many of the cognitive deficits of normal ageing (forgetfulness, distractibility, inflexibility and impaired executive functions) involve prefrontal cortex (PFC) dysfunction<sup>1–4</sup>. The PFC guides behaviour and thought using working memory<sup>5</sup>, which are essential functions in the information age. Many PFC neurons hold information in working memory through excitatory networks that can maintain persistent neuronal firing in the absence of external stimulation<sup>6</sup>. This fragile process is highly dependent on the neurochemical environment<sup>7</sup>. For example, elevated cyclic-AMP signalling reduces persistent firing by opening HCN and KCNQ potassium channels<sup>8,9</sup>. It is not known if molecular changes associated with normal ageing alter the physiological properties of PFC neurons during working memory, as there have been no *in vivo* recordings, to our knowledge, from PFC neurons of aged monkeys. Here we characterize the first recordings of this kind, revealing a marked loss of PFC persistent firing with advancing age that can be rescued by restoring an optimal neurochemical environment. Recordings showed an age-related decline in the firing rate of DELAY neurons, whereas the firing of CUE neurons remained unchanged with age. The memory-related firing of aged DELAY neurons was partially restored to more youthful levels by inhibiting cAMP signalling, or by blocking HCN or KCNQ channels. These findings reveal the cellular basis of age-related cognitive decline in dorsolateral PFC, and demonstrate that physiological integrity can be rescued by addressing the molecular needs of PFC circuits.

Our society is rapidly ageing, with the number of seniors in the United States expected to double by 2050 (United States census, <http://www.census.gov/population/www/pop-profile/elderpop.html>). At the same time, the information age requires increasing organizational skills to deal with even basic needs such as medical care and paying bills. However, executive and working memory functions decline early in the normal ageing process<sup>10–13</sup>, beginning in middle age<sup>14,15</sup>. Thus, cognitive changes with advancing age may be costly, forcing retirement from demanding careers and jeopardizing the ability to live independently in an increasingly complex society. Ageing monkeys provide an ideal model to reveal the neurobiology of normal ageing, as they have a highly developed PFC, but are not subject to age-related dementias<sup>16</sup>. Thus, one can be certain that cognitive changes are the result of normal ageing and not incipient Alzheimer's disease. Like humans, monkeys begin to develop deficits in executive function as early as middle age<sup>17</sup>. Both aged monkeys<sup>18,19</sup> and humans<sup>3,20</sup> are impaired on working memory tasks that require constant updating of the contents of memory (Supplementary Information), bringing to mind information from longer-term stores (for example, where did I leave my car keys this time?), or keeping in mind a recent event (for example, remembering a new phone number).

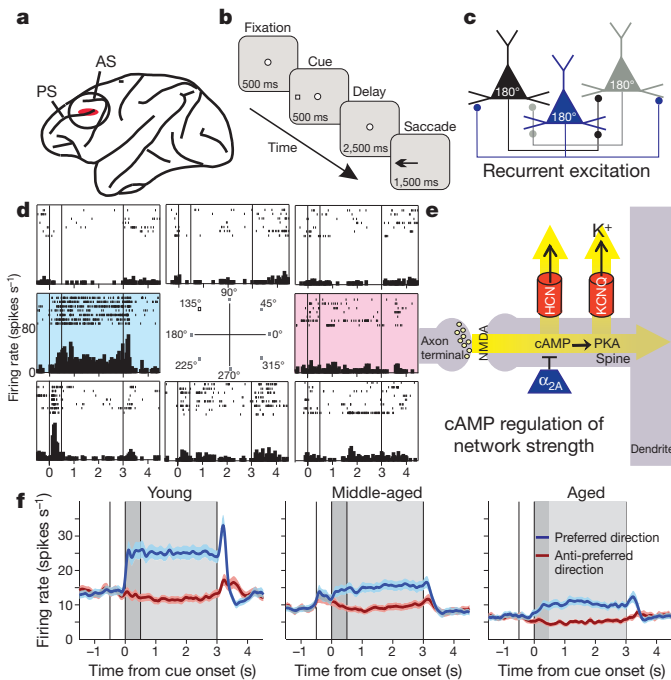
In primates, spatial working memory depends on the highly evolved dorsolateral PFC<sup>6</sup> (Fig. 1a). Spatial working memory performance (Fig. 1b) relies on networks of pyramidal neurons that interconnect at dendritic spines (Fig. 1c), and excite each other to keep information 'in mind', that is, generating persistent spiking activity over a delay

period in a working memory task<sup>6</sup> (Fig. 1d). This ability to maintain information that is no longer in the environment is a fundamental process needed for abstract thought and flexible responding<sup>6</sup>. Intracellular signalling pathways modulate the physiological strength of these recurrent, excitatory PFC network connections<sup>9</sup>. Recent data show that increased cAMP signalling weakens network connectivity by opening potassium channels, whereas inhibiting cAMP signalling and/or closing these channels strengthens connectivity and cognitive ability<sup>9</sup> (Fig. 1e). Specifically, cAMP signalling seems to weaken persistent firing and impair working memory by increasing the open state of HCN (hyperpolarization-activated cyclic nucleotide-gated) channels that are localized on spines where networks interconnect<sup>8</sup>. Recent data indicate that HCN channels may also gate synaptic inputs through interactions with KCNQ channels, whose open state is increased by cAMP-activating protein kinase A (PKA)<sup>21</sup>. Studies indicate that cAMP signalling is disinhibited in the aged PFC<sup>22</sup>. Noradrenergic  $\alpha_{2A}$  receptor inhibition of cAMP may be reduced from loss of  $\alpha_{2A}$  receptors in the aged PFC<sup>23</sup>, and decreased excitation of noradrenergic neurons<sup>24</sup>.

There have been few electrophysiological recordings from aged PFC neurons owing to the demanding nature of this procedure. Recordings from rat orbital PFC found reduced flexibility in aged neurons<sup>25</sup>. However, there have been no *in vivo* recordings from the aged dorsolateral PFC, even though behavioural data indicate that this region is particularly vulnerable to normal ageing. *In vitro* recordings from dorsolateral PFC neurons found relatively subtle changes in excitability with advancing age<sup>26</sup>, but their consequences for executive function must be observed in a cognitively engaged circuit. Here we perform the first physiological characterization of PFC neuronal response during a working memory task in young adult, middle-aged and aged monkeys.

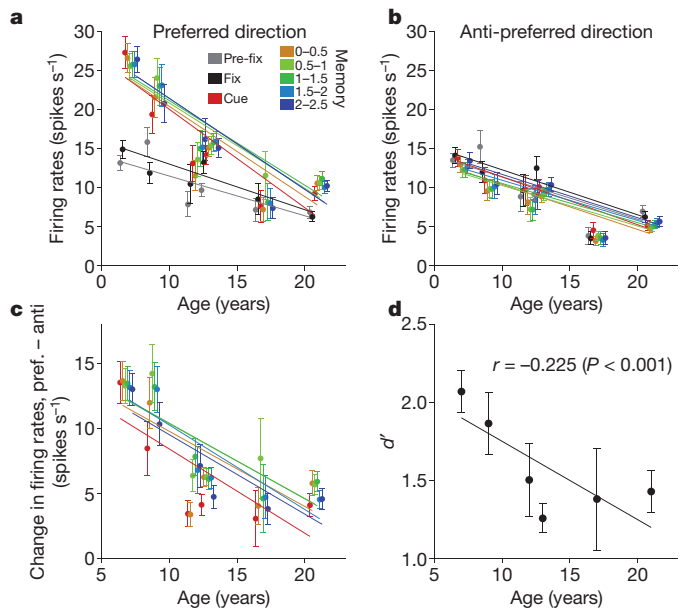
Monkeys (*Macaca mulatta*,  $n = 6$ ) were trained to perform a spatial working memory task in which they have to remember a spatial location over a brief delay period; the spatial location changes randomly on each trial (Fig. 1a). Two animals were young adults (7- and 9-year-old males), two were middle-aged (12- and 13-year-old males), and two were aged (17-year-old male, 21-year-old female). Short delays (2.5 s) were used in all age groups to ensure similar performances (>85% per cent correct) across age groups. Neurons ( $n = 301$ ) were recorded from area 46, the dorsolateral PFC subregion most needed for visuospatial working memory (Fig. 1a). Neurons were characterized based on task-related firing as responsive during (1) the visuospatial cue period, (2) the delay period when the spatial position was being remembered, and/or (3) the motor response period. Some neurons fired only during cue presentation (CUE cells,  $n = 28$ ), whereas most neurons fired during the delay period as well as to the cue and/or response periods (DELAY cells,  $n = 273$ ). Persistent firing during the delay period is of particular interest, as it is required for working memory<sup>6</sup>. Many PFC DELAY neurons increased their activity during the memory of one spatial location (its preferred direction), but not other locations (the 'anti-preferred' direction, 180° away from the preferred direction; Fig. 1d).

<sup>1</sup>Department of Neurobiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA. <sup>2</sup>The John B. Pierce Laboratory, New Haven, Connecticut 06510, USA.



**Figure 1 | Age-related changes in the PFC networks that subserve working memory.** **a**, The region of the dorsolateral PFC most needed for spatial working memory, and the site of recordings. AS, arcuate sulcus; PS, principal sulcus. **b**, The oculomotor delayed response (ODR) spatial working memory task. The monkey fixates on the central stimulus and maintains fixation for the duration of the trial. A cue is briefly presented in one of eight locations, followed by a delay period (2.5 s) in which no spatial information is present. At the end of the delay period, the fixation spot disappears, and the monkey makes an eye movement (saccade) to the remembered location for a juice reward. The cue position randomly changes on subsequent trials. **c**, A diagram of the recurrent excitatory networks subserving working memory. Pyramidal cells with similar spatial tuning excite each other to maintain persistent firing across the delay period<sup>6</sup>. These networks are concentrated in deep layer III<sup>6</sup>. Spatial tuning is enhanced by GABAergic lateral inhibition (not shown). **d**, An example of a dorsolateral PFC DELAY neuron with spatially tuned, persistent firing during the delay period. This neuron shows increased firing for the cue, delay and response for the neuron's preferred direction (highlighted in blue), but not for nonpreferred directions (white backgrounds). The anti-preferred direction opposite to the neuron's preferred direction is shown in red; note that subsequent figures show only the preferred and anti-preferred directions for the sake of brevity. **e**, Pyramidal cells synapse on spines where cAMP–PKA signalling regulates the open state of HCN and KCNQ channels, and thus modulates the strength of network connections<sup>9</sup>. **f**, Population average activity for the dorsolateral PFC DELAY neurons recorded in each age group (102, 101 and 70 neurons for young, middle-aged and old monkeys, respectively). Colours indicate the activity during the trials in which the cue was presented in the neuron's preferred (blue) and anti-preferred (red) directions; the darker grey background refers to the cue period; the lighter grey background to the delay period. Error envelope represents s.e.m.

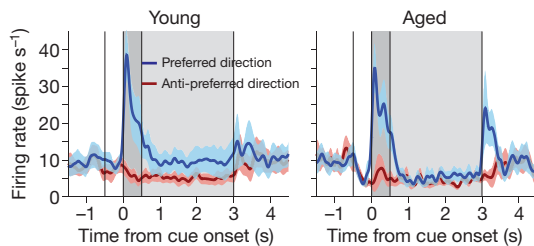
The firing of DELAY cells was markedly reduced with advancing age (Figs 1f and 2). Figure 1f portrays the differences in firing rates across the population of DELAY neurons in young, middle-aged and aged animals (for individual examples of DELAY neurons in young, middle-aged and aged monkeys, see Supplementary Fig. 1). There was a significant decline in the spontaneous firing rate of DELAY cells, as well as a marked decline in task-related firing. Figure 2 shows a steep decline in the firing rates of DELAY cells across the age span ( $t$ -test on age variable in regression analysis,  $P < 10^{-4}$  for all epochs), with older animals showing a restricted range of lower firing rates (Supplementary Fig. 2). This age-related activity decline persisted throughout the 2.5-s delay period (without main effect of epoch (0.5 s) or age  $\times$  epoch interaction in repeated measures ANOVA,  $P > 0.25$ ). Additional control analyses showed that age-related decline in the firing rate of



**Figure 2 | Age-dependent decline in the spatially tuned, persistent firing of dorsolateral PFC DELAY neurons.** **a**, Marked reduction of dorsolateral PFC DELAY activity for the neurons' preferred direction with advancing age. Activity of individual neurons of each animal was averaged separately for the last 0.5 s during the intertrial interval (pre-fix, grey), the fixation period (fix, black), the cue period (cue, red), and the delay period for the neuron's preferred direction. Firing during the delay period is represented in a successive series of 0.5 s intervals (colour coded from yellow to blue). Lines were obtained using linear regression. **b**, Firing rates during the delay period for the anti-preferred direction of the same neurons shown in **a**. There was a significant age-related decline in all epochs, but it was less prominent than the decline in firing for the preferred direction during the delay period. Colour coding for each 0.5 s interval is as in **a**. **c**, Age-related decline in spatial tuning, whereby the difference between firing for the preferred versus anti-preferred directions during the delay period declines with advancing age. Colour coding is as in **a**. **d**, Age-related decline in  $d'$ , that is, the ability to distinguish preferred from anti-preferred spatial directions based on firing rate patterns during the entire delay period. Error bars represent s.e.m.

DELAY cells is not due to a sampling bias during the recording experiment (see Supplementary Information). The age-related decline in firing rate was particularly prominent during the cue and delay periods for the neuron's preferred direction (Fig. 2a); the decline in firing for the anti-preferred direction (Fig. 2b) or before target onset (Fig. 2a) was less pronounced. Consequently, the difference in delay-related firing for the neuron's preferred direction versus its anti-preferred direction eroded with increasing age ( $t$ -test,  $P < 10^{-5}$ , for cue period and every 0.5 s epoch in the delay period; Fig. 2c), largely due to reduced firing for the neuron's preferred direction (Fig. 2a). This led to a reduction in  $d'$  with advancing age ( $t$ -test on age versus  $d'$  correlation coefficient,  $P < 0.0001$ ), that is, a reduced ability to distinguish the preferred from anti-preferred directions during the delay period when spatial information was held in working memory (Fig. 2d). These results are consistent with studies showing impairment in spatial working memory in aged monkeys at relatively short (for example, 5 s) delays<sup>18</sup>, and single-unit data as well as neural circuit modelling indicate that inadequate PFC recurrent network firing underlies the deficits in PFC cognitive function observed in ageing monkeys and humans (Supplementary Figs 3, 4 and Supplementary Information).

In contrast to DELAY neurons, which showed prominent decline in firing with advancing age, there were no age-related changes in the firing rates of PFC CUE cells that responded specifically to the spatial cue (Fig. 3). The average firing rate of these neurons for the preferred direction during the cue period was  $26.7 \pm 4.4$  spikes  $s^{-1}$  in young monkeys ( $n = 12$  neurons), and  $25.3 \pm 3.7$  spikes  $s^{-1}$  in old monkeys (11 neurons;  $t$ -test,  $P > 0.8$ ; significant age  $\times$  cell-type interaction in



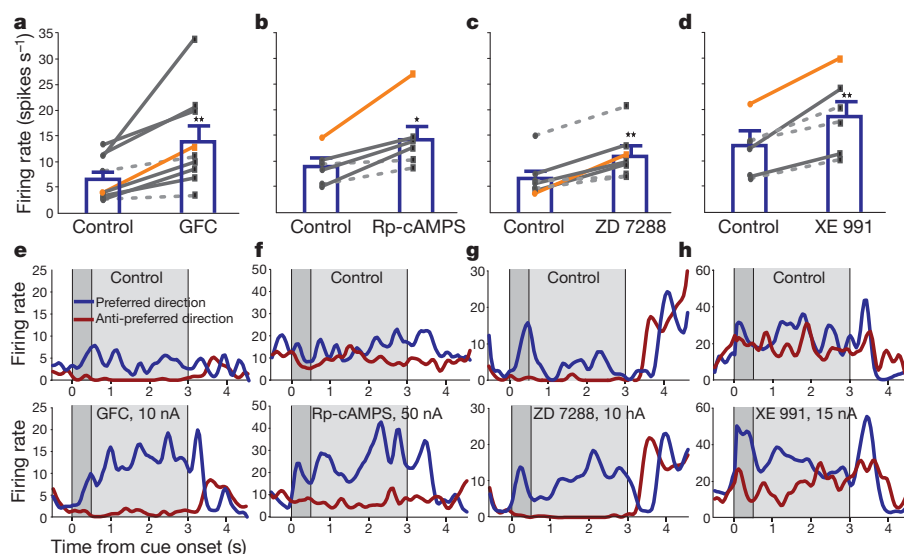
**Figure 3 | Firing rates of dorsolateral PFC CUE cells remain stable in aged monkeys.** The average firing rates of CUE cells in young monkeys (left graph; 10 neurons from a 7-year-old and 2 neurons from a 9-year-old monkey), did not differ from the firing rate in the oldest monkey (right graph; 11 neurons from a 21-year-old monkey), or from the firing rate averaged for both middle-aged (5 neurons from 13-year-old monkey, not shown separately) and old monkeys ( $t$ -test,  $P > 0.7$ ). CUE cells may receive direct, 'bottom-up' excitation from parietal association cortex<sup>6</sup>, which may be less vulnerable to subtle molecular changes with advancing age. A subset of CUE cells recorded in the aged monkey showed high firing rates during the response period. Error envelope represents s.e.m.

two-way ANOVA,  $P < 0.05$ ). These data indicate that reductions in memory-related firing rate do not arise from generalized changes with advancing age affecting all neurons but, rather, are especially evident in recurrent circuits that must maintain firing in the absence of 'bottom-up' sensory stimulation.

It is important to identify which changes in the ageing brain contribute to reduced firing during the delay period. There are many brain alterations associated with advancing age<sup>27</sup>, including decreased PFC grey matter volume<sup>28</sup>, focal changes in white matter<sup>29</sup> and dendritic spine loss<sup>30</sup>, all of which correlate with cognitive decline. Importantly, spine loss is especially prominent in layer III—the layer where the recurrent excitatory networks reside—and thin-type spines are the most vulnerable in the aged PFC<sup>30</sup>. Immunoelectron microscopy indicates that thin spines have the greatest concentration of cAMP–HCN-channel signalling proteins, indicating that disinhibition of cAMP

signalling with advancing age may weaken thin spines in particular<sup>9</sup>. Thus, we tested whether inhibition of cAMP signalling in the PFC could partially restore the working-memory-related firing of aged neurons, or whether reductions in firing were irreversible owing to immutable architectural changes in the aged brain. Drugs were applied near the recorded neurons using iontophoresis, whereby a small electrical current is applied to extrude charged molecules from glass pipettes attached to the recording electrode. Only a minute amount of drug was released, sufficient to alter the firing of nearby neurons, without altering behavioural performance.

Agents that inhibit cAMP signalling or block HCN or KCNQ channels restored persistent firing during the delay period of the working memory task (Fig. 4). For example, iontophoresis of the  $\alpha_{2A}$  agonist, guanfacine (Fig. 4a, e), or the cAMP–PKA inhibitor, Rp-cAMPS (Fig. 4b, f), significantly increased firing during the delay period on trials when the cue had appeared at the neuron's preferred direction. In contrast, the PDE4 inhibitor, etazolate—which increases cAMP signalling—further decreased neuronal firing in aged neurons ( $P < 0.001$ ; Supplementary Fig. 5). We also tested whether blockade of HCN or KCNQ channels could restore firing, given that cAMP–PKA signalling increases the open state of these ion channels. As shown in Fig. 4c and g, a low dose of the HCN channel blocker ZD7288 significantly enhanced the delay-related firing rates of neurons in aged monkeys. KCNQ channels were also of interest, as *in vitro* physiological characterizations of PFC neurons in aged primates have found increases in the slow afterhyperpolarization, which is mediated in part by KCNQ channels<sup>26</sup>. As shown in Fig. 4d and h, blockade of KCNQ channels with XE991 increased delay-related firing in aged PFC neurons. Thus, agents that reduced cAMP opening of HCN or KCNQ channels all restored firing rates to levels resembling those observed in younger monkeys. These findings are consistent with behavioural data showing that guanfacine and Rp-cAMPS can enhance working memory performance in aged animals when administered systemically (guanfacine) or directly into the rat PFC (guanfacine or Rp-cAMPS)<sup>22</sup> (see Supplementary Information). On the basis of these data, guanfacine



**Figure 4 | Iontophoresis of compounds that inhibit cAMP–PKA signalling, or block HCN or KCNQ channel signalling, strengthens delay-related firing in aged PFC DELAY neurons.** a–d, A summary of the results showing a significant increase in population-average firing rate for the neuron's preferred direction compared to control conditions (paired  $t$ -test,  $P < 0.01$  for a, c and d, and  $P < 0.05$  for b) following iontophoresis of the  $\alpha_{2A}$  adrenergic agonist guanfacine applied at 10 nA (GFC; a, significant effects in 7 out of 9 neurons,  $t$ -test,  $P < 0.05$ , indicated by solid lines); the cAMP inhibitor Rp-cAMPS at 50 nA (b, significant in 4 out of 6 neurons); the HCN channel blocker, ZD7288 at 15 nA (c, significant in 4 out of 7 neurons); and the KCNQ channel blocker

XE991 at 15 nA (d, significant in 3 out of 6 neurons). In all cases, significant effects were found more frequently than expected by chance (binomial test,  $P < 0.005$ ). The orange lines represent the individual neurons shown in e–h. e–h, Individual examples of neurons under control conditions (top) firing to their preferred (blue trace) or anti-preferred (red trace) directions, compared to their firing patterns following iontophoresis of guanfacine (e), Rp-cAMPS (f), ZD7288 (g) or XE991 (h). The orange lines in a–d indicate the individual neurons shown in e–h. Error bars are s.e.m. \* $P < 0.05$ , \*\* $P < 0.01$  significant difference between drug versus control for the neuron's preferred direction.



is currently being tested in elderly humans with PFC cognitive deficits (<http://www.clinicaltrials.gov>: trial number NCT00935493).

The current study revealed a physiological basis for age-related working memory decline in the primate brain, with a reduction of memory-related firing beginning in middle age and worsening with advancing age. This marked change in network physiology may render higher cortical circuits especially vulnerable to neurodegenerative processes such as Alzheimer's disease. However, these studies also uncovered more hopeful data showing that restitution of the proper neurochemical environment can partially restore physiological integrity. These data establish that cognitive changes with advancing age are malleable, and that there is potential to restore at least some cognitive abilities in the elderly. Maintaining strong PFC physiology into advanced age will be an important advantage in an increasingly complex, ageing society.

## METHODS SUMMARY

All experiments were performed in accordance with National Institutes for Health guidelines for animal research and were approved by the Yale Institutional Animal Care and Use Committee.

**Behavioural training.** Monkeys were trained on the oculomotor delayed response (ODR) task (Fig. 1b) as reported elsewhere<sup>8</sup>, with special care to minimize all stress. The aged monkeys took longer to learn the task, rested more frequently than young monkeys during the testing sessions, and performed fewer trials each day than younger animals. A brief delay period (2.5 s) was chosen to ensure that all monkeys in the study performed above 85% correct during training, and could maintain high levels of performance during the study (Supplementary Fig. 6).

**Physiology and iontophoresis.** Single unit recordings were made from the dorsolateral PFC surrounding the caudal portion of the principal sulcus (Fig. 1a). The recording procedures and iontophoresis methods have been described previously<sup>8</sup>. Single neuron activity was readily isolated (Supplementary Fig. 7). Denoting the average activity during the 0.5 s cue period and that during the 2.5 s delay period during the trials in preferred directions as *C* and *D*, neurons were classified as CUE cells, when the ratio *D/C* was >0.5, and as DELAY cells otherwise. Statistical analyses, including ANOVA and regression analyses, were performed using Matlab (MathWorks). The effect of age on firing rate during a series of 0.5 s epochs was tested using a regression model, and its statistical significance with a *t*-test.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 9 December 2010; accepted 23 May 2011.

Published online 27 July 2011.

- West, R. L. An application of prefrontal cortex function theory to cognitive aging. *Psychol. Bull.* **120**, 272–292 (1996).
- Cabeza, R., Anderson, N. D., Houle, S., Mangels, J. A. & Nyberg, L. Age-related differences in neural activity during item and temporal-order memory retrieval: a positron emission tomography study. *J. Cogn. Neurosci.* **12**, 197–206 (2000).
- Gazzaley, A., Cooney, J. W., Rissman, J. & D'Esposito, M. Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neurosci.* **8**, 1298–1300 (2005).
- Prakash, R. S. *et al.* Age-related differences in the involvement of the prefrontal cortex in attentional control. *Brain Cogn.* **71**, 328–335 (2009).
- Goldman-Rakic, P. S. in *Handbook of Physiology, The Nervous System, Higher Functions of the Brain* Vol. 5 (ed. Plum, F.) 373–417 (American Physiological Society, 1987).
- Goldman-Rakic, P. S. Cellular basis of working memory. *Neuron* **14**, 477–485 (1995).
- Robbins, T. W. & Arnsten, A. F. The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* **32**, 267–287 (2009).
- Wang, M. *et al.*  $\alpha$ 2A-adrenoceptor stimulation strengthens working memory networks by inhibiting cAMP-HCN channel signaling in prefrontal cortex. *Cell* **129**, 397–410 (2007).

- Arnsten, A. F. T., Paspalas, C. D., Gamo, N. J., Yang, Y. & Wang, M. Dynamic network connectivity: a new form of neuroplasticity. *Trends Cogn. Sci.* **14**, 365–375 (2010).
- Bowles, R. P. & Salthouse, T. A. Assessing the age-related effects of proactive interference on working memory tasks using the Rasch model. *Psychol. Aging* **18**, 608–615 (2003).
- Royall, D. R., Palmer, R., Chiodo, L. K. & Polk, M. J. Normal rates of cognitive change in successful aging: the freedom house study. *J. Int. Neuropsychol. Soc.* **11**, 899–909 (2005).
- Burke, S. N. & Barnes, C. A. Neural plasticity in the ageing brain. *Nature Rev. Neurosci.* **7**, 30–40 (2006).
- Cappell, K. A., Gmeindl, L. & Reuter-Lorenz, P. A. Age differences in prefrontal recruitment during verbal working memory maintenance depend on memory load. *Cortex* **46**, 462–473 (2010).
- Davis, H. P. *et al.* Lexical priming deficits as a function of age. *Behav. Neurosci.* **104**, 288–297 (1990).
- Bucur, B. & Madden, D. J. Effects of adult age and blood pressure on executive function and speed of processing. *Exp. Aging Res.* **36**, 153–168 (2010).
- Sisodia, S. S., Martin, L. J., Walker, L. C., Borchelt, D. R. & Price, D. L. Cellular and molecular biology of Alzheimer's disease and animal models. *Neuroimaging Clin. N. Am.* **5**, 59–68 (1995).
- Moore, T. L., Killiany, R. J., Herndon, J. G., Rosene, D. L. & Moss, M. B. Executive system dysfunction occurs as early as middle-age in the rhesus monkey. *Neurobiol. Aging* **27**, 1484–1493 (2006).
- Rapp, P. R. & Amaral, D. G. Evidence for task-dependent memory dysfunction in the aged monkey. *J. Neurosci.* **9**, 3568–3576 (1989).
- Herndon, J. G., Moss, M. B., Rosene, D. L. & Killiany, R. J. Patterns of cognitive decline in aged rhesus monkeys. *Behav. Brain Res.* **87**, 25–34 (1997).
- Rypma, B. & D'Esposito, M. Isolating the neural mechanisms of age-related changes in human working memory. *Nature Neurosci.* **3**, 509–515 (2000).
- George, M. S., Abbott, L. F. & Siegelbaum, S. A. Hyperpolarization-activated HCN channels inhibit subthreshold EPSPs through voltage-dependent interactions with M-type  $K^+$  channels. *Nature Neurosci.* **12**, 577–584 (2009).
- Ramos, B. *et al.* Dysregulation of protein kinase A signaling in the aged prefrontal cortex: new strategy for treating age-related cognitive decline. *Neuron* **40**, 835–845 (2003).
- Moore, T. L. *et al.* Cognitive impairment in aged rhesus monkeys associated with monoamine receptors in the prefrontal cortex. *Behav. Brain Res.* **160**, 208–221 (2005).
- Downs, J. L. *et al.* Orexin neuronal changes in the locus coeruleus of the aging rhesus macaque. *Neurobiol. Aging* **28**, 1286–1295 (2007).
- Schoenbaum, G., Setlow, B., Saddoris, M. P. & Gallagher, M. Encoding changes in orbitofrontal cortex in reversal-impaired aged rats. *J. Neurophysiol.* **95**, 1509–1517 (2006).
- Luebke, J. I. & Chang, Y. M. Effects of aging on the electrophysiological properties of layer 5 pyramidal cells in the monkey prefrontal cortex. *Neuroscience* **150**, 556–562 (2007).
- Luebke, J., Barbas, H. & Peters, A. Effects of normal aging on prefrontal area 46 in the rhesus monkey. *Brain Res. Rev.* **62**, 212–232 (2010).
- Alexander, G. E. *et al.* Age-related regional network of magnetic resonance imaging gray matter in the rhesus macaque. *J. Neurosci.* **28**, 2710–2718 (2008).
- Peters, A. *et al.* Neurobiological bases of age-related cognitive decline in the rhesus monkey. *J. Neuropathol. Exp. Neurol.* **55**, 861–874 (1996).
- Dumitriu, D. *et al.* Selective changes in thin spine density and morphology in monkey prefrontal cortex correlate with aging-related cognitive impairment. *J. Neurosci.* **30**, 7507–7515 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This research was supported by PHS grant P01AG030004 from the National Institute on Aging. The authors would like to thank J. Thomas, L. Ciavarella, S. Johnson, B. Brunson and M. Horn for their assistance in making this work possible.

**Author Contributions** M.W. and X.-J.W., J.A.M., D.L. and A.F.T.A. designed the experiments. M.W. carried out all the physiology experiments, with the help of Y.Y., N.J.G., L.E.J. and J.A.M. Data analyses were performed by M.W., D.L., J.A.M. and M.L. Computational modelling was performed by X.-J.W. All authors participated in the writing of the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interest: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to A.F.T.A. ([amy.arnsten@yale.edu](mailto:amy.arnsten@yale.edu)).

## METHODS

**Oculomotor delayed response task.** Studies were performed on four adult male rhesus monkeys (*Macaca mulatta*) trained on the spatial oculomotor delayed response (ODR) task as previously described<sup>8</sup>. This task requires the monkey to make a memory-guided saccade to a remembered visuospatial target. Each trial began when the subject fixated at the central spot for 0.5 s (fixation period). Subsequently, a cue was illuminated for 0.5 s at one of eight peripheral targets (cue period). After the cue was extinguished, a 2.5 s delay period followed. During the cue and delay periods, the monkey was required to maintain central fixation. At the end of the delay, the fixation spot was extinguished, instructing the monkey to make a memory-guided saccade to the previously cued location (saccade period). The monkey was rewarded with fruit juice immediately after every successful response. The position of the stimulus was randomized over trials such that it had to be remembered on a trial-by-trial basis. The subject's eye position was monitored with the ISCAN Eye Movement Monitoring System, and the ODR task was generated by the TEMPO real-time system (Reflective Computing). The aged monkeys took longer to learn the task, rested more frequently than young monkeys during the testing sessions, and performed fewer trials each day than younger animals. A brief delay period (2.5 s) was chosen to ensure that all monkeys in the study performed above 85% correct during training, and could maintain high levels of performance during the study (Supplementary Fig. 6).

**Recording locus.** Before recording, the animal underwent a magnetic resonance image (MRI) scan to obtain exact anatomical coordinates of brain structures, which guided the placement of the chronic recording chambers. MRI-compatible materials were used for the implant so that the position of the recording chambers could be confirmed by MRI after implantation. The site of recordings in the present study were located in an area ranging from 0–5 mm anterior to the caudal end of the principal sulcus and –2–2 mm medial to the principal sulcus.

**Pharmacology, physiology and data acquisition.** Guanfacine, XE991 and ZD7288 (Tocris) and etazolate (Sigma) were dissolved at 0.01 M in triple-distilled water (adjusted with HCl to pH 3.5–4.0). Rp-cAMP (Sigma) was dissolved at 0.01 M in triple-distilled water (adjusted with NaOH to pH 9).

Iontophoretic electrodes were constructed with a 20- $\mu$ m pitch carbon fibre (ELSI) inserted in the central barrel of a seven-barrel non-filamented capillary glass (Friedrich and Dimmock). The assembly was pulled using a multipipette electrode puller (MicroData Instrument) and the tip was bevelled to obtain the finished electrode. Finished electrodes had impedances of 0.3–1.0 M $\Omega$  (at 1 kHz) and tip sizes of 30–40  $\mu$ m. The outer barrels of the electrode were then filled with 3 drug solutions (two consecutive barrels each) and the solutions were pushed to the

tip of the electrode using compressed air. A Neurophore BH2 iontophoretic system (Medical Systems) was used to control of the delivery of the drugs. The drug was ejected at currents that varied from 5–50 nA. Retaining currents of –3 to –5 nA were used in a cyclical manner (1 s on, 1 s off) when not applying drugs. Drug ejection did not create noise in the recording, and there was no systematic change in either spike amplitude or time course at any ejection current.

The electrode was mounted on a MO-95 micromanipulator (Narishige) in a 25-gauge stainless steel guide tube. The dura was punctured using the guide tube to facilitate access of the electrode to cortex. Extracellular voltage was amplified using an AC/DC differential amplifier (A-M systems; Model 3000) and band-pass filtered (180 Hz–6 KHz, 20 dB gain, 4-pole Butterworth; Kron-Hite). Signals were digitized (20.83 kHz, micro 1401, Cambridge Electronics Design) and acquired using the Spike2 software (CED). Neural activity was analysed using waveform sorting by a template-matching algorithm, which made it possible to isolate more than one unit at the same recording site. Post-stimulus time histograms and rastergrams were constructed online to determine the relationship of unit activity to the task. Unit activity was measured in spikes per second. If the rastergrams displayed task-related activity, the units were recorded further and pharmacological testing was performed.

Data were first collected from the cell under a control condition in which at least eight trials at each of eight cue locations were obtained. On establishing the stability of the cells' activity, this control condition was followed by the drug application. Dose-dependent effects of the drug were tested in two or more consecutive conditions. Drugs were continuously applied at a relevant current throughout a given condition. Each condition had ~8 (typically 10 or more) trials at each location for statistical analyses of effects.

**Data analysis.** For purposes of data analysis, each trial in the ODR task was divided into four epochs: fixation, cue, delay and response (saccade). The fixation epoch lasted for 0.5 s. The cue epoch lasted for 0.5 s and corresponds to the stimulus presentation phase of the task. Delay lasted for 2.5 s and reflects the mnemonic component of the task. The response phase started immediately after the delay epoch and lasted ~1.5 s.

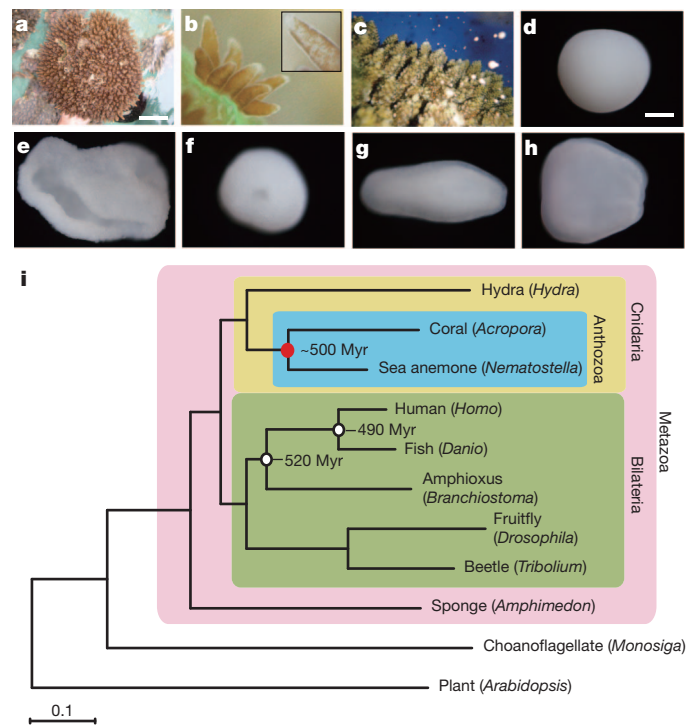
Two-way analysis of variance, ANOVA, was used to examine the spatial tuned task-related activity with regard to (1) different periods of the task (cue, delay, response versus fixation); and (2) different cue locations. One-way ANOVAs were used to assess the effect of the drug application on cells showing delay-related activity. Statistical analyses, including ANOVA and regression analyses were performed using Matlab (MathWorks). The effect of age on firing rate during a series of 0.5 s epochs was tested using a regression model, and its statistical significance with a *t*-test.

# Using the *Acropora digitifera* genome to understand coral responses to environmental change

Chuya Shinzato<sup>1\*</sup>, Eiichi Shoguchi<sup>1\*</sup>, Takeshi Kawashima<sup>1\*</sup>, Mayuko Hamada<sup>1\*</sup>, Kanako Hisata<sup>1</sup>, Makiko Tanaka<sup>1</sup>, Manabu Fujie<sup>2</sup>, Mayuki Fujiwara<sup>1</sup>, Ryo Koyanagi<sup>1</sup>, Tetsuro Ikuta<sup>1</sup>, Asao Fujiyama<sup>3</sup>, David J. Miller<sup>4</sup> & Nori Satoh<sup>1</sup>

Despite the enormous ecological and economic importance of coral reefs, the keystone organisms in their establishment, the scleractinian corals, increasingly face a range of anthropogenic challenges including ocean acidification and seawater temperature rise<sup>1–4</sup>. To understand better the molecular mechanisms underlying coral biology, here we decoded the approximately 420-megabase genome of *Acropora digitifera* using next-generation sequencing technology. This genome contains approximately 23,700 gene models. Molecular phylogenetics indicate that the coral and the sea anemone *Nematostella vectensis* diverged approximately 500 million years ago, considerably earlier than the time over which modern corals are represented in the fossil record (~240 million years ago)<sup>5</sup>. Despite the long evolutionary history of the endosymbiosis, no evidence was found for horizontal transfer of genes from symbiont to host. However, unlike several other corals, *Acropora* seems to lack an enzyme essential for cysteine biosynthesis, implying dependency of this coral on its symbionts for this amino acid. Corals inhabit environments where they are frequently exposed to high levels of solar radiation, and analysis of the *Acropora* genome data indicates that the coral host can independently carry out *de novo* synthesis of mycosporine-like amino acids, which are potent ultra-violet-protective compounds. In addition, the coral innate immunity repertoire is notably more complex than that of the sea anemone, indicating that some of these genes may have roles in symbiosis or coloniality. A number of genes with putative roles in calcification were identified, and several of these are restricted to corals. The coral genome provides a platform for understanding the molecular basis of symbiosis and responses to environmental changes.

Coral reefs are estimated to harbour around one third of all described marine species<sup>6</sup>, and their productivity supports around one quarter of marine fisheries, but declines in coral abundance and wholesale loss of reef habitats are one of the most pressing environmental issues of our time. The major architects of coral reefs, the scleractinian corals, are anthozoan cnidarians that form obligate endosymbioses with photosynthetic dinoflagellates of the genus *Symbiodinium* (Fig. 1b). The symbionts confer on the coral holobiont the ability to fix CO<sub>2</sub> and to deposit the massive aragonite (a form of calcium carbonate) skeletons that distinguish reef-building corals from other anthozoans such as sea anemones. The association is fragile, however, and collapses under stress. Despite the ecological and economic significance of corals, the molecular mechanisms underlying much of coral biology—including stress responses and disease—remain unclear, but it is clear that corals retain much of the complex gene repertoire of the ancestral metazoan<sup>7</sup>. To address the lack of molecular data for reef-building corals, we determined the whole-genome sequence of *A. digitifera* (Fig. 1a–h), a dominant species on Okinawan reefs. Not only are *Acropora* species the dominant reef-building corals of the Indo-Pacific, but they are also among the most sensitive of corals to increased seawater temperatures<sup>8</sup>.



**Figure 1 | The coral *Acropora digitifera* and an early occurrence of corals on Earth.** **a**, The colony the genome of which was sequenced in the present study. This colony is maintained in aquarium culture at the Sesoko Station, University of the Ryukyus, and is thus available for further investigation of the genome. Scale bar, 10 cm. **b**, Polyps of the coral showing the presence of symbiotic dinoflagellates (*Symbiodinium* sp.) (inset, enlargement). **c**, Natural spawning of the coral. **d–h**, Eggs, embryos, larva and primary polyp of *A. digitifera*, from which messenger RNA was extracted for transcriptome analyses. Scale bar, 200  $\mu$ m. **i**, Molecular phylogeny of corals. 94,200 aligned amino acid positions of proteins encoded by 422 genes were obtained from the sponge *Amphimedon queenslandica*, from the cnidarians *A. digitifera*, *Nematostella vectensis* and *Hydra magnipapillata*, and from the triploblasts *Tribolium castaneum*, *Drosophila melanogaster*, *Branchiostoma floridae*, *Danio rerio* and *Homo sapiens*. The sequences were analysed using maximum likelihood methods, with the plant *Arabidopsis thaliana* and the choanoflagellate *Monosiga brevicollis* serving as outgroups. The scale bar represents 0.1 expected substitutions per site in the aligned regions. The topology was supported by 100% bootstrap value. Approximate divergent times of the occurrence of basal chordates and divergence of vertebrates lineages are shown. This analysis indicates a deeper divergence of *Acropora* and *Nematostella*, approximately 500 million years (Myr) ago.

<sup>1</sup>Marine Genomics Unit, Okinawa Institute of Science and Technology Promotion Corporation, Onna, Okinawa 904-0412, Japan. <sup>2</sup>DNA Sequencing Center Section, Okinawa Institute of Science and Technology Promotion Corporation, Onna, Okinawa 904-0412, Japan. <sup>3</sup>National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan. <sup>4</sup>ARC Centre of Excellence for Coral Reef Studies and School of Pharmacy and Molecular Sciences, James Cook University, Townsville, Queensland 4811, Australia.

\*These authors contributed equally to this work.



On the basis of flow cytometry, the *A. digitifera* genome is approximately 420 Mbp (Supplementary Figs 1 and 2) and is therefore similar in size to that of the sea anemone *Nematostella*<sup>9</sup>. Sperm from a single colony served as the source of DNA for sequencing using a combination of Roche 454 GS-FLX<sup>10</sup> and Illumina Genome Analyser IIx (GAIIx)<sup>11</sup> methods. The genome was sequenced to approximately 151-fold coverage (Supplementary Table 1), enabling the generation of an assembly comprising a total of 419 Mbp (Supplementary Tables 2–5; contig N50 = 10.7 kbp and scaffold N50 = 191.5 kbp; Supplementary Fig. 3). The genome is approximately 39% G+C (Supplementary Fig. 4), and contains 23,668 predicted protein-coding loci (Supplementary Table 6). Transposable elements occupy approximately 12.9% of the genome (Supplementary Table 7). The coral gene set is comparable in size and composition with those of *Nematostella vectensis*<sup>9</sup> and *Hydra magnipapillata*<sup>12</sup> (Supplementary Tables 6, 8 and 9). The genome browser is accessible at [http://marinegenomics.oist.jp/acropora\\_digitifera](http://marinegenomics.oist.jp/acropora_digitifera) (Supplementary Fig. 5). Approximately 93% of the *A. digitifera* genes have matches in other metazoans (Supplementary Fig. 6a), and of these, 11% have clear homology only among expressed sequence tag (EST) data from corals<sup>13</sup> (Supplementary Fig. 6b), suggesting the presence of a considerable number of coral-specific genes.

Corals are morphologically very similar to sea anemones, but their evolutionary origins are obscure. Reef-building Scleractinia first appeared in the fossil record in the mid-Triassic (approximately 240 million years ago)<sup>5</sup>, but were already highly diversified, suggesting much earlier origins. The availability of fully sequenced genomes for three cnidarians—*Acropora* (the present study), *Nematostella*<sup>9</sup> and *Hydra*<sup>12</sup>—allowed the estimation of the depth of the divergence between corals and other metazoans. Molecular phylogenetic analyses based on an alignment of 94,200 amino acid positions suggest a divergence time of *Acropora* and *Nematostella* between 520 to ~490 million years ago (the late Cambrian or early Ordovician) (Fig. 1i). The implied earlier origin of Scleractinia indicates that corals have persisted through previous periods of major environmental change, including the mass extinction event at the Permian/Triassic boundary, when global CO<sub>2</sub> and temperature were much higher than at present. However, whereas the Scleractinia as a lineage has persisted on evolutionary time scales, whether modern coral reefs can adapt to rapid environmental change on ecological time scales is a very different question.

The obligate endosymbiosis of corals dates at least from the mid-Triassic, and the longevity of this association might therefore be expected to have resulted in changes within the coral genome. We were unable to find any *Symbiodinium* DNA sequences in the coral genome, hence there is as yet no evidence for horizontal gene transfer from symbiont to host (Supplementary Fig. 6). However, comparative analyses indicated that, in the case of *Acropora*, the coral host might be metabolically dependent on the symbiont. Using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database<sup>14</sup>, the metabolic repertoire of *Acropora* was compared to that of its non-symbiotic relative, the sea anemone *Nematostella* (Supplementary Table 10), leading to the identification of an apparent metabolic deficiency in *Acropora*. The biosynthesis of cysteine from homocysteine and/or serine requires the activities of two enzymes, cystathionine β-synthase (Cbs) and cystathionase (cystathionine γ-lyase; Cth) (Table 1). Whereas we were able to identify genes encoding the latter in both *A. digitifera* and *Nematostella*, the former could not be identified in *Acropora* despite a clear match being present in *Nematostella* (Supplementary Fig. 7). Although extensive transcriptomic data are available for various *Acropora* spp<sup>13</sup>, we could find no evidence for a *Cbs* transcript in any of these. Moreover, whereas a polymerase chain reaction (PCR) strategy confirmed the presence of *Cbs* in some other corals, *Galaxea fascicularis*, *Favites chinensis*, *Favia lizardensis* and *Ctenactis echinata*, no amplification products could be obtained for two different *Acropora* species (Table 1 and Supplementary Fig. 8). Although the analyses presented here do not rigorously exclude the presence of *Cbs* activity in *Acropora*,

**Table 1 | The presence or absence of a gene encoding Cbs for L-cysteine biosynthesis**

	L-Homocysteine + L-Serine	Cbs	L-Cystathionine	Cth	L-Cysteine
<i>Hydra magnipapillata</i>		Yes*		Yes	
<i>Nematostella vectensis</i>		Yes*		Yes	
Complexa					
<i>Acropora digitifera</i>		No†		Yes	
<i>Acropora tenuis</i>		No‡		ND	
<i>Acropora millepora</i>		No§		Yes	
<i>Galaxea fascicularis</i>		Yes‡		ND	
Robusta					
<i>Montastraea faveolata</i>		Yes§		Yes	
<i>Favia lizardensis</i>		Yes‡		ND	
<i>Favites chinensis</i>		Yes‡		ND	
<i>Ctenactis echinata</i>		Yes‡		ND	

ND, not determined.

\* Supported by sequenced genome and EST analyses.

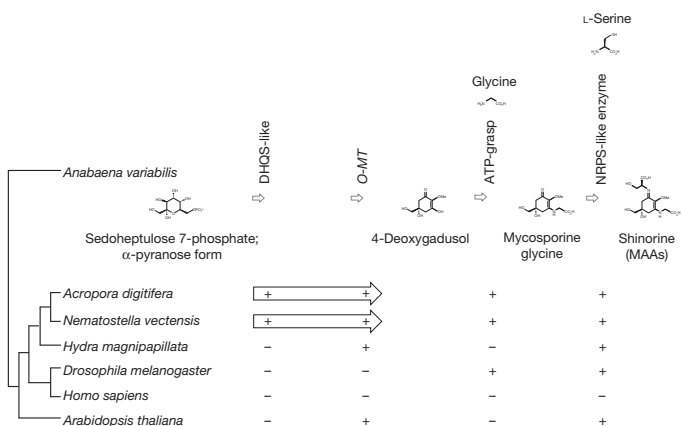
† Supported by sequenced genome, EST and PCR amplification of genomic DNA.

‡ Supported by PCR amplification of genomic DNA.

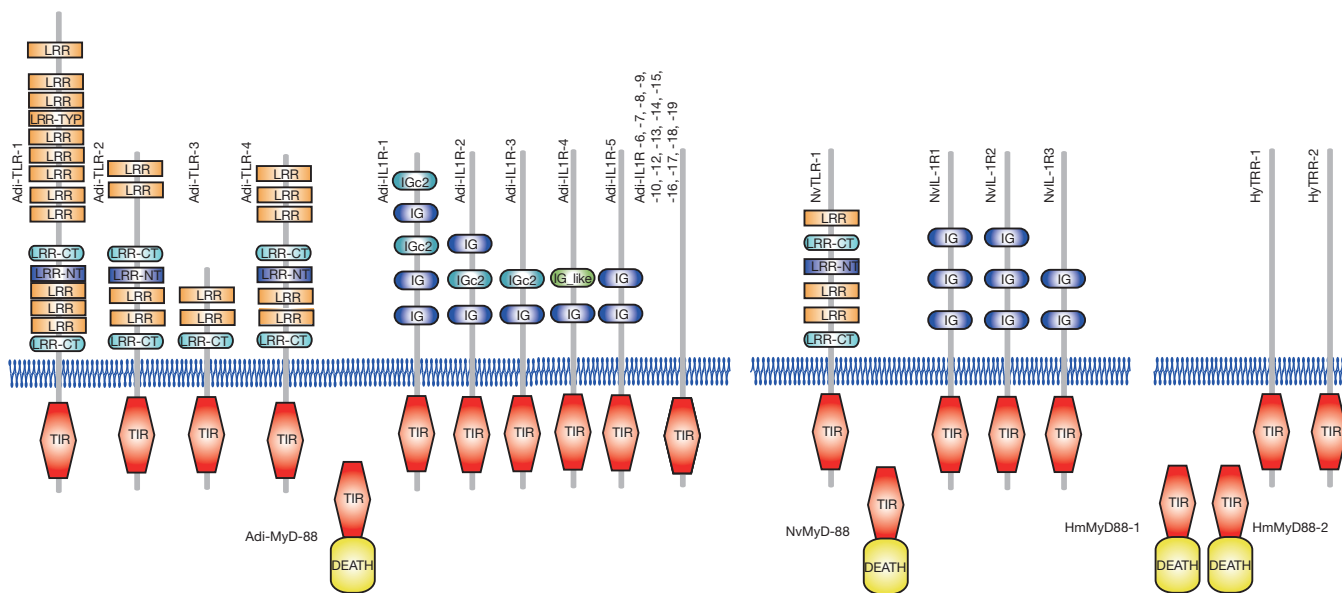
§ Supported by EST analyses.

they raise the intriguing possibility of a metabolic basis for the obligate nature of symbiosis in *Acropora*; differences in dependency could potentially explain not only the phenomenon of symbiont selectivity, but also the high sensitivity of *Acropora* to environmental challenges.

Reef-building corals typically inhabit shallow and relatively clear tropical waters and are therefore constantly exposed to high levels of ultraviolet irradiation. As corals are particularly susceptible to bleaching when exposed to both raised temperatures and high solar radiation<sup>2,4</sup>, one intriguing question is how corals protect themselves against ultraviolet damage. Photo-protective compounds, such as the mycosporine amino acids (MAAs), have been isolated from corals<sup>15,16</sup> but, because similar compounds have been identified in algae, the sources of these compounds were unknown. Recently a short (four-step) pathway encoded by a gene cluster (DHQS-like, O-MT, ATP-grasp and NRPS-like) (Fig. 2 and Supplementary Figs 9–12) has been demonstrated to be both necessary and sufficient in the cyanobacterium *Anabaena variabilis* to convert pentose-phosphate metabolites to shinorine, a photo-protective MAA<sup>17</sup>. Scanning the available whole-genome data allowed us to identify clear homologues of all four members of the cyanobacterial shinorine gene cluster in both *A. digitifera* and *N. vectensis* (Fig. 2), indicating that both the coral and the sea anemone have the ability to carry out *de novo* synthesis of ultraviolet-protective compounds. Hence, MAA synthesis in corals and other cnidarians is not symbiont dependent.



**Figure 2 | The genes required for the biosynthesis of shinorine are present in anthozoan cnidarians.** Top, the organization of the gene cluster involved in the biosynthetic pathway of the photo-protective molecule shinorine, a mycosporine-like amino acid, in the cyanobacterium *Anabaena variabilis*. Bottom, the presence of corresponding genes in various organisms is indicated (+). The *Acropora* and *Nematostella* genomes contain homologues of each of the four genes, in which DHQS-like and O-MT are fused with each other.



**Figure 3 | Repertoires of TIR-domain-containing proteins of three cnidarians.** The schematic representation of the domain structures of TIR-domain-containing proteins identified in *A. digitifera*, alongside the corresponding complements from *Nematostella vectensis* and *Hydra*

Surveys of *Acropora* for genes associated with innate immunity<sup>18</sup>, apoptosis<sup>19</sup> and autophagy<sup>19</sup> indicate not only the complexity of these systems in *Acropora* (Supplementary Figs 13–23), but also that the coral innate immune repertoire is more sophisticated than that of *Nematostella*. For example, whereas a single canonical Toll/TLR protein is present in *N. vectensis*<sup>18</sup>, the *Acropora* genome encodes at least four such molecules, as well as five IL-1R-related proteins and a number of TIR-only proteins (Fig. 3). Likewise, the *Acropora* repertoire of NACHT/NB-ARC domains, which are characteristic of primary intracellular pattern receptors<sup>20</sup>, is again highly complex: an order of magnitude more NACHT/NB-ARC domains are present in coral than in other animals (Supplementary Table 11), and some of these cnidarian proteins have novel domain structures (Supplementary Fig. 23b). In terms of the apparent expansion and divergence of NACHT-encoding genes, the coral resembles amphioxus<sup>21</sup>, the sea urchin<sup>22</sup> and angiosperms<sup>23</sup>. The greater complexity of the coral innate immunity network may in part reflect adaptations associated with the symbiotic state and coloniality.

The coral repertoire of genes with predicted roles in skeleton deposition is of particular interest given the likely impact of ocean acidification resulting from rising atmospheric CO<sub>2</sub> on coral calcification. Surveys of the *Acropora* genome for specific groups of proteins associated with calcification, including the eukaryotic-type carbonic anhydrases<sup>24</sup> are given in Supplementary Table 12. In general, the soluble fraction of the organic matrix in scleractinian corals is very rich in acidic amino acids, and has a particularly high aspartic acid composition<sup>25</sup>. A number of candidate organic-matrix proteins were identified in *Acropora* (Supplementary Fig. 24). For several of these, orthologues could be identified in *A. millepora* and/or *A. palmata* but only one of these (Adi-SAP6) was found in other coral species (Supplementary Table 13). Galaxins, first purified from the coral *Galaxea fascicularis*, are unique to corals and are the only coral skeletal matrix protein for which the complete primary structure has been determined<sup>26</sup>. However, galaxin possesses neither acidic regions (the fraction of Asp+Asn in the galaxin is 9.7%) nor obvious Ca<sup>2+</sup>-binding domains<sup>26</sup>. Four genes encoding galaxin-related proteins were

*magnipapillata*. The repertoire of Toll/TLR, IL-1R-like and TIR-only proteins is significantly more complex in the case of *A. digitifera* than in *N. vectensis* or *H. magnipapillata*. TIR, TIR domain. DEATH, DEATH domain. IG and IGc2, Ig domain. LRR, LRR-TRY, LRR-CT and LRR-NT, leucine-rich repeats.

identified in the *A. digitifera* genome (Supplementary Fig. 25), including two likely *A. digitifera* homologues of Gfa-galaxin.

Here we decoded the 420-Mbp genome of the reef-building coral *Acropora digitifera*, with the aim of providing a platform for understanding the molecular basis of symbiosis and responses to environmental change. Some of the main findings are: (1) a relatively deep divergence of the lineage leading to the reef-building corals; (2) although we could find no evidence for horizontal gene transfer from symbiont to coral despite the long evolutionary history of the association, *Acropora* may have lost a gene essential for cysteine biosynthesis and thus be metabolically dependent on its symbionts; (3) the coral host has the ability to independently carry out *de novo* synthesis of the MAA family of photoprotective compounds; (4) the innate immune repertoire of coral is highly complex in comparison with the non-symbiotic and solitary sea anemone *Nematostella*; and (5) a number of coral-specific gene families are likely to have evolved in the context of calcification. These data also provide a basis for systems biology approaches to understanding the establishment, function and collapse of coral symbioses. If and when a whole-genome sequence becomes available for the dinoflagellate symbiont of corals *Symbiodinium* sp. (zooxanthellae), these resources will together provide additional perspectives on the symbiosis and a powerful resource for understanding the response of the holobiont to environmental stresses such as raised seawater temperatures or ocean acidification.

## METHODS SUMMARY

Sperm DNA obtained from a single colony of the coral *Acropora digitifera* was used for genome sequencing by Roche 454 GS-FLX<sup>10</sup> and Illumina Genome Analyser Ix (GAIIx)<sup>11</sup>. The 454 shotgun and paired-end reads were assembled *de novo* by GS De novo Assembler version 2.3 (Newbler, Roche)<sup>10</sup>, and subsequent scaffolding was performed by SOPRA<sup>27</sup> and SPACEx<sup>28</sup> using the Illumina mate-pair information. Transcriptome analysis was also performed. A set of gene model predictions (the *A. digitifera* Gene Model v. 1) was generated mainly by AUGUSTUS<sup>29</sup>, and a genome browser has been established using the Generic Genome Browser (GBrowser) 2.17. The annotation and identification of *Acropora* genes were performed by three approaches, individual methods or combinations of the methods:

reciprocal BLAST analyses, screening the gene models against the Pfam database<sup>30</sup> and phylogenetic analyses.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 7 January; accepted 3 June 2011.**

**Published online 24 July 2011.**

- Hughes, T. P. *et al.* Climate change, human impacts, and the resilience of coral reefs. *Science* **301**, 929–933 (2003).
- Hoegh-Guldberg, O. *et al.* Coral reefs under rapid climate change and ocean acidification. *Science* **318**, 1737–1742 (2007).
- Carpenter, K. E. *et al.* One-third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science* **321**, 560–563 (2008).
- Weis, V. M. Cellular mechanisms of cnidarian bleaching: stress causes the collapse of symbiosis. *J. Exp. Biol.* **211**, 3059–3066 (2008).
- Stanley, G. D. Jr & Fautin, D. G. Paleontology and evolution. The origins of modern corals. *Science* **291**, 1913–1914 (2001).
- Wilkinson, C. *Status of Coral Reefs of the World* (Australian Institute of Marine Studies, 2004).
- Miller, D. J., Ball, E. E. & Technau, U. Cnidarians and ancestral genetic complexity in the animal kingdom. *Trends Genet.* **21**, 536–539 (2005).
- Loya, Y. *et al.* Coral bleaching: the winners and the losers. *Ecol. Lett.* **4**, 122–131 (2001).
- Putnam, N. H. *et al.* Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* **317**, 86–94 (2007).
- Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380 (2005).
- Bentley, D. R. Whole-genome re-sequencing. *Curr. Opin. Genet. Dev.* **16**, 545–552 (2006).
- Chapman, J. A. *et al.* The dynamic genome of *Hydra*. *Nature* **464**, 592–596 (2010).
- Hemrich, G. & Bosch, T. C. Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem cell differentiation. *Bioessays* **30**, 1010–1018 (2008).
- Kanehisa, M. *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36**, D480–D484 (2008).
- Shick, J. M. & Dunlap, W. C. Mycosporine-like amino acids and related Gadusols: biosynthesis, accumulation, and UV-protective functions in aquatic organisms. *Annu. Rev. Physiol.* **64**, 223–262 (2002).
- Rastogi, R. P. *et al.* Photoprotective compounds from marine organisms. *J. Ind. Microbiol. Biotechnol.* **37**, 537–558 (2010).
- Balskus, E. P. & Walsh, C. T. The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria. *Science* **329**, 1653–1656 (2010).
- Miller, D. J. *et al.* The innate immune repertoire in cnidaria—ancestral complexity and stochastic gene loss. *Genome Biol.* **8**, R59 (2007).
- Dunn, S. R., Schnitzler, C. E. & Weis, V. M. Apoptosis and autophagy as mechanisms of dinoflagellate symbiont release during cnidarian bleaching: every which way you lose. *Proc. R. Soc. Lond. B* **274**, 3079–3085 (2007).
- Meylan, E., Tschopp, J. & Karin, M. Intracellular pattern recognition receptors in the host response. *Nature* **442**, 39–44 (2006).
- Huang, S. *et al.* Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Res.* **18**, 1112–1126 (2008).
- Hibino, T. *et al.* The immune gene repertoire encoded in the purple sea urchin genome. *Dev. Biol.* **300**, 349–365 (2006).
- Leister, D. Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. *Trends Genet.* **20**, 116–122 (2004).
- Jackson, D. J., Macis, L., Reithner, J., Degnan, B. M. & Wörheide, G. Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis. *Science* **316**, 1893–1895 (2007).
- Sarashina, I. & Endo, K. Skeletal matrix proteins of invertebrate animals: comparative analysis of their amino acid sequences. *Paleontological Res.* **10**, 311–336 (2006).
- Fukuda, I. *et al.* Molecular cloning of a cDNA encoding a soluble protein in the coral exoskeleton. *Biochem. Biophys. Res. Commun.* **304**, 11–17 (2003).
- Dayarian, A., Michael, T. P. & Sengupta, A. M. SOPRA: scaffolding algorithm for paired reads via statistical optimization. *BMC Bioinformatics* **11**, 345 (2010).
- Boetzer, M. *et al.* Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578–579 (2011).
- Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics* **24**, 637–644 (2008).
- Finn, R. D. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **38**, D211–D222 (2010).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We acknowledge A. Iguchi, K. Sakai and all other staff members of the Sesoko Station, Tropical Biosphere Research Center, University of the Ryukyus for their help in collection and maintenance of the coral, M. Hidaka and V. Weis for suggestions, Y. Yokoyama, T. Usami and other staff members of our Unit for sequencing, and S. Brenner, R. Baughman and T. Ichikawa for their support. D. Rokhsar and J. Chapman are acknowledged for suggestions on sequence assembly and gene prediction. The super-computing was supported by the IT Section of OIST and the Human Genome Center, University of Tokyo. This work was supported in part by Grants-in-Aids from MEXT and JSPS, Japan.

**Author Contributions** Project design and coordination: N.S., C.S., E.S., T.K., M.H. and D.J.M. Genome and EST sequencing and assembly: C.S., T.K., E.S., K.H., M. Fujie, M. Fujiwara, M.T., M.H., A.F. and R.K. Annotation and analysis: E.S., C.S., M.H., T.K., K.H., M.T., R.K. and T.I. Writing: N.S., C.S., E.S., T.K., M.H. and D.J.M.

**Author Information** The genome assembly has been deposited with the DNA Databank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL) and GenBank under project accession BACK01000001–BACK01053640 (contigs) and DF093604–DF097774 (scaffolds). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike license, and is freely available to all readers at [www.nature.com/nature](http://www.nature.com/nature). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to N.S. ([norisky@oist.jp](mailto:norisky@oist.jp)).



## METHODS

**Biological specimen.** Under permits from the Aquaculture Agency of Okinawa Prefecture (the number 20–27), part of an *A. digitifera* colony was collected and has subsequently been maintained in an aquarium at the Sesoko Station, Tropical Biosphere Research Center, University of the Ryukyus.

**The number of chromosomes, diploidy and genome size of *Acropora digitifera*.**

The number of chromosomes was determined by their preparation from nuclei of embryonic cells. The diploidy of the genome was examined by fluorescent *in situ* hybridization (FISH) of BAC clones<sup>31</sup>, which were constructed in pKS146 (ref. 32). The genome size was estimated by flow cytometry<sup>33</sup> using sperm nuclei from the same colony that was used to sequence the genome.

**Genome sequencing and assembly.** The sperm was obtained from the single colony and sperm DNA was used for genome sequencing and BAC library construction. Genome sequence data were obtained using single read, paired-end and mate-pair protocols on the Roche 454 GS-FLX<sup>10</sup> and Illumina GAIIx<sup>11</sup> instruments. The genomic DNA was fragmented, libraries prepared and sequencing conducted according to the manufacturer's protocols. The 454 shotgun and paired-end reads were assembled *de novo* by GS *De novo* Assembler version 2.3 (Newbler, Roche)<sup>10</sup> in heterozygotic mode with adjusted algorithms to reflect an increase in the expected variability in sequence identity. Possible PCR duplicates in Illumina mate-pair reads were removed by MarkDuplicates in Picard tools (<http://picard.sourceforge.net>), and then subsequent scaffolding of the 29,765 Newbler output was performed by SOPRA<sup>27</sup> and SSPACE<sup>28</sup> using the Illumina mate-pair information. Gaps inside the scaffolds were closed with Illumina paired-end data using GapCloser<sup>34</sup>. To overcome potential assembly errors arising from tandem repeats, sequences that were aligned to another sequence over 50% of the length by BLASTN ( $1 \times 10^{-50}$ ) were removed from the assembly<sup>35</sup>.

**Transcriptome analyses.** RNA was isolated from eggs, gastrulae, planulae, polyps and adults. Total RNA was extracted following the manufacturer's instructions (Invitrogen) and purified using DNase and an RNeasy micro kit (QIAGEN). Transcriptome libraries for 454 GS-FLX were prepared<sup>36</sup> and sequenced as per manufacturer's instructions. In addition, Illumina 50-bp paired-end RNA-seq sequencing was performed. All high-quality sequences (quality value  $\geq 15$ ) were assembled by a Velvet/Oases assembler<sup>37</sup> with hash length 27.

**Gene prediction.** A set of gene model predictions (the *A. digitifera* Gene Model v. 1) was generated using AUGUSTUS<sup>29</sup>. AUGUSTUS 2.0.4 was trained on the 877 EST assemblies recommended by PASA<sup>38</sup> for this purpose. The gene models were created by running AUGUSTUS on a repeat-masked genome produced by RepeatMasker<sup>39</sup>, and improved by PASA<sup>38</sup>. A genome browser has been established using the assembled genome sequences using the Generic Genome Browser (GBrowser) 2.17 (ref. 40).

**Identification of *Acropora* genes involved in the response to environmental change.** Three approaches, individual methods or combinations of the methods,

were used to annotate the protein-coding genes in the *A. digitifera* genome. A primary approach to the identification of putative orthologues of *A. digitifera* genes was reciprocal BLAST analysis. This was carried out on the basis of mutual best hit in BLAST analyses for human, mouse, or *Drosophila* genes against the *A. digitifera* gene models (BLASTP) or the assembly (BLASTN). A second approach used in the case of genes encoding proteins with one or more specific protein domains, was to screen the merged models against the Pfam database (Pfam-A.hmm, release 24.0; <http://pfam.sanger.ac.uk>)<sup>30</sup>, which contains 11,912 conserved domains using HMMER (hmmer3)<sup>41</sup>. In the case of complex multigene families, a third annotation method was used; sets of related sequences were subjected to phylogenetic analyses to determine more precisely orthology relationships. For these purposes, amino acid sequences were aligned using ClustalW<sup>42</sup> or ClustalX<sup>42</sup> under the default options. Gaps and ambiguous areas were excluded using Gblocks 0.91b<sup>43</sup> with the default parameters and then checked manually. On the basis of the alignment data sets, phylogenetic trees were constructed by neighbour joining and/or maximum likelihood. Calculations and tree construction were performed in SeaView<sup>44</sup>. The KEGG pathway database<sup>14</sup> was used to examine the metabolic repertoire of *Acropora* in comparison to that of the sea anemone *Nematostella*.

31. Shoguchi, E. *et al.* Chromosomal mapping of 170 BAC clones in the ascidian *Ciona intestinalis*. *Genome Res.* **16**, 297–303 (2006).
32. Fujiyama, A. *et al.* Construction and analysis of a human-chimpanzee comparative clone map. *Science* **295**, 131–134 (2002).
33. Davies, D. & Allen, P. in *Flow Cytometry: Principles and Applications* (ed. Macey, M. G.) (Humana, 2007).
34. Li, R. *et al.* The sequence and *de novo* assembly of the giant panda genome. *Nature* **463**, 311–317 (2010).
35. Wurm, Y. *et al.* The genome of the fire ant *Solenopsis invicta*. *Proc. Natl Acad. Sci. USA* **108**, 5679–5684 (2011).
36. Meyer, E. *et al.* Sequencing and *de novo* analysis of a coral larval transcriptome using 454 GSFLX. *BMC Genomics* **10**, 219 (2009).
37. Zerbino, D. R. & Birney, E. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).
38. Haas, B. J. *et al.* Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654–5666 (2003).
39. Smit, A. F. A., Hubley, R. & Green, P. RepeatMasker Open-3.0. (<http://www.repeatmasker.org>) (1996–2010).
40. Stein, L. D. *et al.* The generic genome browser: a building block for a model organism system database. *Genome Res.* **12**, 1599–1610 (2002).
41. Eddy, S. R. Profile hidden Markov models. *Bioinformatics* **14**, 755–763 (1998).
42. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
43. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).
44. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224 (2010).

# Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching

Marcelo J. Berardi<sup>1,2</sup>, William M. Shih<sup>2,3,4</sup>, Stephen C. Harrison<sup>1,2,5</sup> & James J. Chou<sup>1,2</sup>

Mitochondrial uncoupling protein 2 (UCP2) is an integral membrane protein in the mitochondrial anion carrier protein family, the members of which facilitate the transport of small molecules across the mitochondrial inner membrane<sup>1,2</sup>. When the mitochondrial respiratory complex pumps protons from the mitochondrial matrix to the intermembrane space, it builds up an electrochemical potential<sup>2</sup>. A fraction of this electrochemical potential is dissipated as heat, in a process involving leakage of protons back to the matrix<sup>2</sup>. This leakage, or ‘uncoupling’ of the proton electrochemical potential, is mediated primarily by uncoupling proteins<sup>2</sup>. However, the mechanism of UCP-mediated proton translocation across the lipid bilayer is unknown. Here we describe a solution-NMR method for structural characterization of UCP2. The method, which overcomes some of the challenges associated with membrane-protein structure determination<sup>3</sup>, combines orientation restraints derived from NMR residual dipolar couplings (RDCs) and semiquantitative distance restraints from paramagnetic relaxation enhancement (PRE) measurements. The local and secondary structures of the protein were determined by piecing together molecular fragments from the Protein Data Bank that best fit experimental RDCs from samples weakly aligned in a DNA nanotube liquid crystal. The RDCs also determine the relative orientation of the secondary structural segments, and the PRE restraints provide their spatial arrangement in the tertiary fold. UCP2 closely resembles the bovine ADP/ATP carrier (the only carrier protein of known structure<sup>4</sup>), but the relative orientations of the helical segments are different, resulting in a wider opening on the matrix side of the inner membrane. Moreover, the nitroxide-labelled GDP binds inside the channel and seems to be closer to transmembrane helices 1–4. We believe that this biophysical approach can be applied to other membrane proteins and, in particular, to other mitochondrial carriers, not only for structure determination but also to characterize various conformational states of these proteins linked to substrate transport.

The elucidation almost 50 years ago of the mechanism of oxidative phosphorylation provided the first example of the coupling of membrane transport to energy production<sup>5</sup>. The mitochondrial respiratory complex transfers electrons from nutrients to oxygen, pumping protons from the mitochondrial matrix to the intermembrane space as it builds up an electrochemical potential of ~200 mV (Supplementary Fig. 1). Most of this potential is converted to work that drives ATP synthesis, but a fraction is dissipated as heat in the uncoupling of the proton electrochemical potential by members of a family of proteins in the mitochondrial inner membrane, collectively termed uncoupling proteins<sup>2</sup> (UCPs). The proton translocation activity of UCPs requires fatty acids as cofactors; it is inhibited by GDP. UCP1 was first identified in the mitochondria of brown adipose tissue as a protein responsible for thermoregulation mediated by such tissue<sup>6</sup>. Unlike UCP1, the primary function of UCP2 may be related to lowering production of reactive oxygen species<sup>2</sup>. Expression of UCP2 in pancreatic  $\beta$ -cells is strongly correlated with decreased insulin secretion, and UCP2

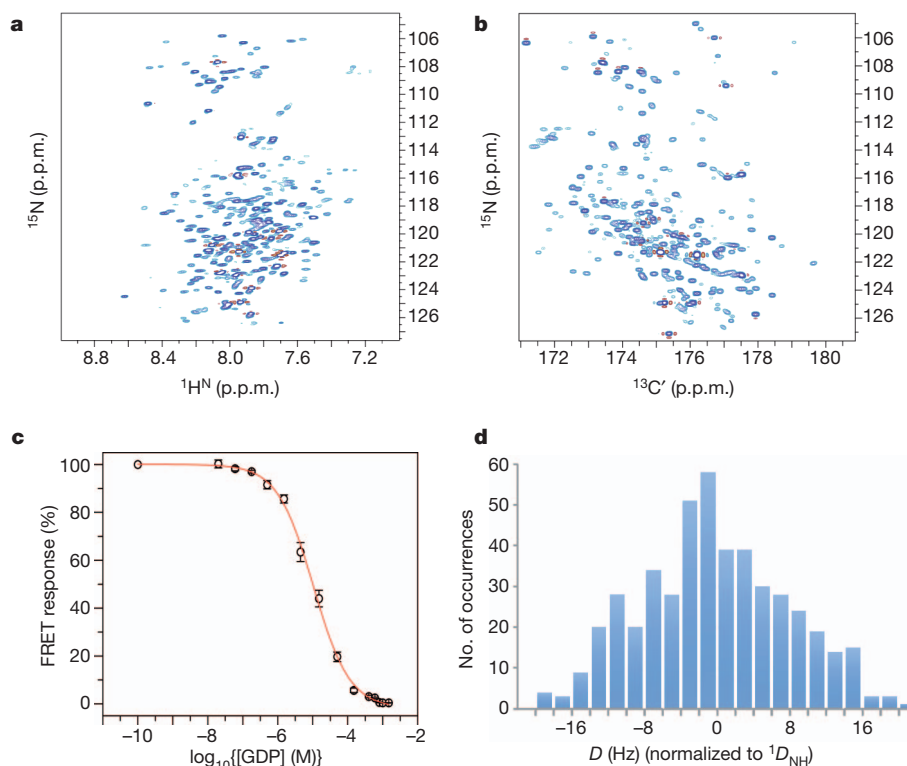
inactivation restores the glucose-sensing pathway<sup>7</sup>. UCP2 also has a role in glucagon secretion from pancreatic  $\alpha$ -cells<sup>8</sup> and in regulation of hunger by means of ghrelin in neurons containing neuropeptide Y and agouti-related protein<sup>9</sup>. More recently, UCP2 has been linked to chemotherapy resistance and metabolic reprogramming in cancer<sup>10,11</sup>.

NMR spectroscopy has yielded the structures of a number of membrane proteins, including channels<sup>12–14</sup>, a redox protein<sup>15</sup>, a kinase<sup>16</sup> and sensory rhodopsin<sup>17</sup>. In many of these studies, a major issue was the purification of the protein and its reconstitution in a model membrane medium supporting protein solubility and stability. In the case of UCP2, reconstitution in dodecylphosphocholine (DPC) micelles with a small amount of lipid (cardiolipin and dimyristoyl phosphatidylcholine (DMPC)) yielded usable NMR spectra (Fig. 1a, b). We prepared homogeneous UCP2 by passing detergent-solubilized protein through a GDP-analogue column, which selects for native protein. Samples used for NMR measurements were at pH 6.5 and contained 5 mM GDP, 150 mM DPC, 1 mM cardiolipin and 2 mM DMPC. Under those conditions, a fluorescence resonance energy transfer (FRET) experiment showed that GDP binds to UCP2 with an apparent dissociation constant of ~5  $\mu$ M (Fig. 1c). When the NMR sample was incorporated into liposomes, it showed fatty-acid-activated and GDP-inhibited proton translocation activity (Supplementary Fig. 2). Negative-stain electron microscopy showed monodisperse, channel-like particles (Supplementary Fig. 3).

Local structure determination by NMR typically involves assigning nuclear Overhauser enhancements between backbone amide protons and aliphatic protons. In the case of UCP2, it was difficult to collect a sufficient number of unambiguous nuclear Overhauser enhancements because of severe overlap of backbone (<sup>1</sup>H, <sup>15</sup>N) resonances. The peaks in the 3D HNCO spectrum are mostly separated, however. The HNCO experiment, which correlates the chemical shifts of backbone <sup>1</sup>H<sup>N</sup>, <sup>15</sup>N and <sup>13</sup>C' nuclides, has been widely used for measuring one-bond RDCs. For structure determination, we therefore relied on orientation restraints derived from RDCs. In a macromolecule marginally oriented in a magnetic field, the RDC between a pair of spin-1/2 nuclides encodes orientations of the internuclear vector connecting the two nuclides. The most effective medium for weakly orienting protein in solution is a liquid crystal formed by large, elongated particles<sup>18</sup>; for membrane proteins, these particles must also resist a high concentration of detergent. We used DNA nanotubes, constructed by ‘DNA origami’, which can form a stable nematic phase unaffected by detergent<sup>19</sup> (Supplementary Fig. 4). The nanotube alignment medium allowed us to measure RDCs for internuclear vectors <sup>1</sup>H<sup>N</sup>–<sup>15</sup>N (<sup>1</sup>D<sub>NH</sub>), <sup>13</sup>C'–<sup>13</sup>C $\alpha$  (<sup>1</sup>D<sub>C'C $\alpha$</sub> ) and <sup>15</sup>N–<sup>13</sup>C' (<sup>1</sup>D<sub>NC'</sub>) (Fig. 1d and Supplementary Fig. 5). On average, there were 2.2 RDCs per residue for regions with confirmed resonance assignment.

With fewer than three RDCs per residue, calculating the local structures using the restrained molecular dynamics protocol presents a severe local minimum problem, because of the intrinsic orientation degeneracy of the dipolar coupling function. We used an approach that

<sup>1</sup>Jack and Eileen Connors Structural Biology Laboratory, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>2</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>3</sup>Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. <sup>4</sup>Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts 02138, USA. <sup>5</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, Massachusetts 02115, USA.



**Figure 1 | NMR spectra, GDP binding and outline of RDC-based molecular fragment assignment.** **a**, The  $^1\text{H}$ - $^{15}\text{N}$  transverse-relaxation-optimized spectroscopy (TROSY) heteronuclear single quantum coherence spectrum of  $^{15}\text{N}$ -,  $^{13}\text{C}$ - and  $^2\text{H}$ -labelled UCP2 reconstituted in DPC micelles (with 5 mM GDP), recorded at pH 6.5, 33 °C and a  $^1\text{H}$  frequency of 600 MHz. **b**, The ( $^{15}\text{N}$ ,  $^{13}\text{C}'$ ) projection of the three-dimensional (3D) TROSY-HNCO spectrum of the sample in **a**, recorded under the same conditions. p.p.m., parts per million.

exhaustively searches the Protein Data Bank for fragments of structure that agree with the experimental data. This method, commonly known as molecular fragment replacement (MFR), was first applied in crystallography to build molecular fragments into crystallographically determined electron density<sup>20</sup>. A later NMR study demonstrated that with about four RDCs per residue, it was possible to determine the backbone structure of ubiquitin using molecular fragments fit to RDCs<sup>21</sup>. More recently, a fragment search method that combines experimentally determined chemical shift and RDCs has been implemented in the CS-ROSETTA structure modelling software and shown to be effective in determining structures of water-soluble proteins<sup>22,23</sup>.

We constructed a database containing 320,000 seven-residue fragments extracted from structures of 1,279 water-soluble and 235 membrane-associated proteins. For each seven-residue stretch along the UCP2 sequence, we fitted the corresponding RDCs to all fragments in the database, using singular value decomposition<sup>24</sup>. We evaluated the quality of fit with the parameter  $Q_{\text{free}}$ , which is a normalized root mean squared difference between experimental RDCs and RDCs predicted independently by the structural model<sup>25</sup>. After this exhaustive search, we collected fragments with  $Q_{\text{free}} < 25\%$  and for which the magnitude and rhombicity of the alignment tensor were within 10% of their true values (Fig. 1d). These ‘candidate fragments’ were used to determine local backbone structure according to a three-step protocol. The first step is ‘fragment assignment’, whereby the candidate fragments are sorted and those with the smallest  $Q_{\text{free}}$  values are assigned to the corresponding protein segments until no more fragments can be placed without overlap with already assigned segments (Fig. 2a). The second step is ‘gap filling’. Here, for any of the gap regions (with fewer than four residues) not assigned in step one, we search for the candidate fragment that overlaps in the backbone dihedral angles  $\phi$  and  $\psi$  with the two flanking fragments and that provides the lowest  $Q_{\text{free}}$

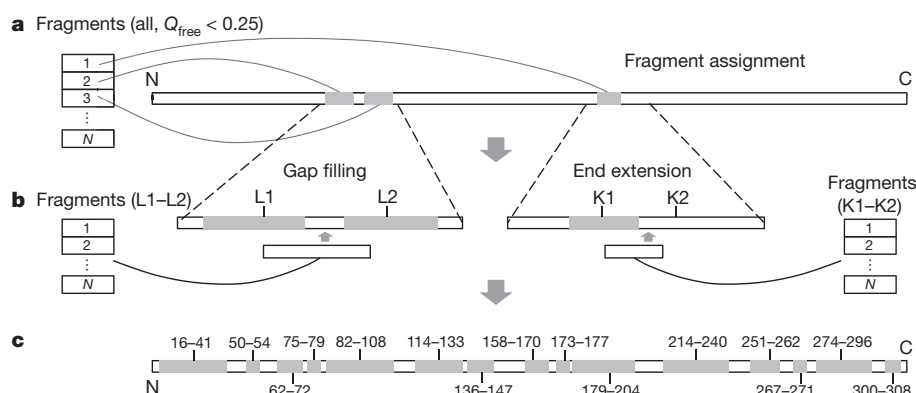
Comparison of **a** and **b** illustrates that most resonances can be resolved in the 3D HNCO spectrum. **c**, Specific binding of GDP to UCP2 under the conditions in **a**. The FRET response (Methods) follows displacement of fluorescent MANT-GDP by GDP as the concentration of the latter increases. [GDP], GDP concentration. **d**, Histogram of 470 unambiguously assigned vectors  $^1D_{\text{NH}}$ ,  $^1D_{\text{C}\alpha}$  or  $^1D_{\text{NC}'}$  (all normalized to  $^1D_{\text{NH}}$ ). The magnitude and rhombicity of the alignment tensors are 10 Hz and 0.61, respectively.

value for the merged segment consisting of the newly filled gap structure and the two original flanking fragments (Fig. 2b). In the third step, ‘end extension’, we extend the amino- or carboxy-terminal end of an assigned segment by searching for a fragment that overlaps best with the original segment in  $\phi$  and  $\psi$ , and that provides the lowest  $Q_{\text{free}}$  value for the final extended segment (Fig. 2b). Steps two and three were repeated until no more structure could be assigned with confidence. Using this procedure, we identified 15 structured segments (Fig. 2c).

In addition to RDCs, we obtained semiquantitative distance restraints from PRE measurements. The paramagnetic moiety introduced for PRE measurement is the nitroxide spin label in MTSL (Methods), which we covalently attached at cysteines. PRE restraints have been demonstrated in low-resolution structure determination of both water-soluble and membrane proteins<sup>16,26</sup>. Four UCP2 samples were produced, each with a spin label attached at a unique position (Cys 68, Cys 105, Cys 202 and Cys 255); they all showed similar NMR spectra (Supplementary Fig. 6). These samples together provided 452 PRE distance restraints (an important subset is shown in Fig. 3a).

Finally, to determine the tertiary structure, we strongly enforced the backbone  $\phi$  and  $\psi$  of the 15 MFR-derived structural segments while applying the PRE and RDC restraints. During the structure calculation, the MFR segments made up most of the local structure of the protein, and the PRE and RDC measurements provided spatial and orientational restraints, respectively, for these structured segments. The calculation generated an ensemble of 15 structures with a backbone root mean squared deviation of 1.3 Å (Fig. 3b). There are no experimental data on the side chains. The structure was determined in the presence of GDP, a well-known inhibitor of UCP2 activity. Although side-chain resonance assignments, which are required to obtain an atomic-resolution view of GDP binding, were not available, we investigated GDP binding qualitatively by the PRE method. We



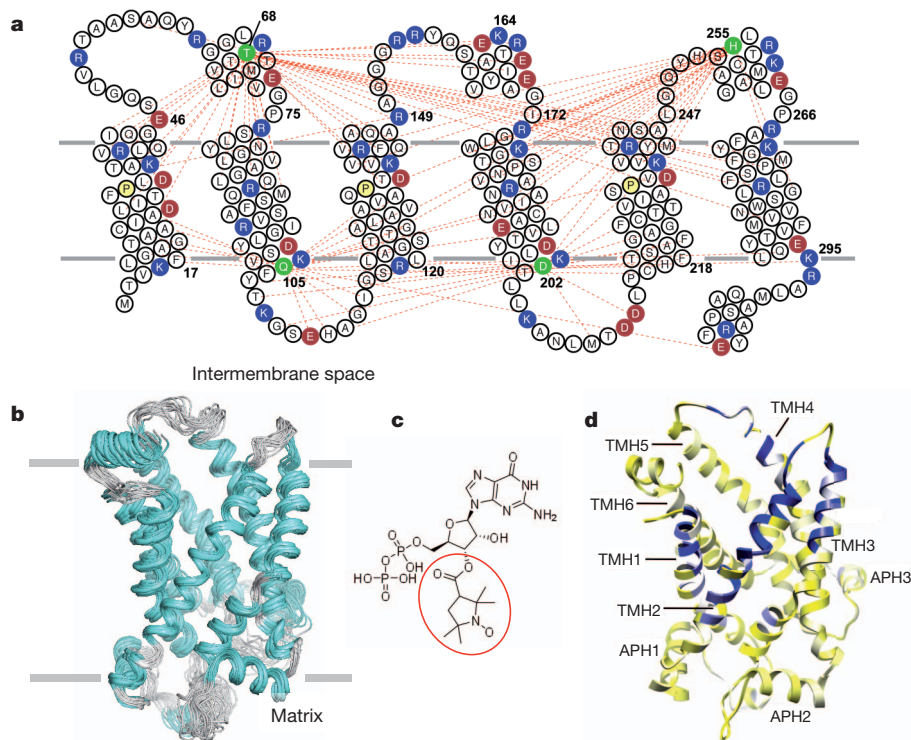


**Figure 2 | Operations involved in RDC-based structural segment building.** **a**, Initial fragment assignment. **b**, Gap filling (left) and end extension (right). **c**, The 15 continuous structured segments of UCP2 determined by RDC-based

MFR (shaded and labelled). Details of these operations are described in the main text and in Methods.

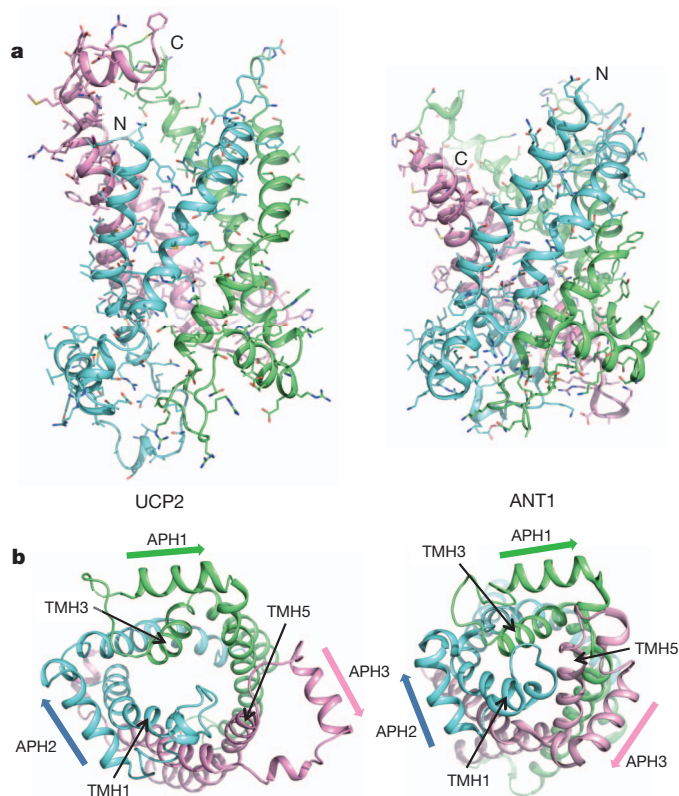
used a paramagnetic nitroxide derivative of GDP (NO-GDP; Fig. 3c) and measured broadening of protein resonances by the spin label. Mapping the magnitude of the PRE to the structure indicates that NO-GDP binds within the UCP2 channel, but the PRE seems to span a large area, covering residues of transmembrane helices (TMHs) 1–4 (Fig. 3d). The distribution of PRE-sensitive residues indicates that the nitroxide radical is in the half of the channel facing the intermembrane space. On the basis of the PRE-derived position of the nitroxide moiety, we located the GDP portion of NO-GDP by maximizing electrostatic interactions between the GDP and charged residues in the channel (Supplementary Fig. 7). The model shows that the PRE data are consistent with GDP binding deep within the UCP2 channel, similar to the proposed ADP-binding site in ANT1<sup>27,28</sup> (also known as SLC25A4).

The overall conformation of UCP2 closely resembles that of the bovine ADP/ATP carrier<sup>4</sup> (ANT1), despite their low (~20%) sequence identity. They are both channel-like structures, in which three pseudo-repeats (Fig. 3a) adopt similar folds (Fig. 4a). Each repeat consists of a TMH (odd numbered), a loop, an amphipathic helix (APH) and a second TMH (even numbered). Moreover, as in ANT1, the TMHs have kinks at prolines that are conserved in the carrier-protein sequences (Fig. 3a). The structural similarity with ANT1 further supports the notion that the members of the large carrier family have a conserved structure and that small variations within the conserved fold govern the specificity of substrate binding and translocation. The most striking differences between the UCP2 and ANT1 structures are in the third repeat. In each of the three repeats of ANT1, the APH packs



**Figure 3 | Solution structure of UCP2 and region of GDP binding.** **a**, UCP2 sequence and membrane topology, with basic and acidic residues shown in blue and red, respectively. The conserved prolines at the proline kinks of TMHs 1, 3 and 5 are shown in yellow. The spin-labelled positions are highlighted in green. The red dashed lines represent long-range or interhelical PRE distances (<19 Å) between the spin-label and backbone amide protons. **b**, An ensemble of 15 low-energy structures derived from NMR restraints. The backbone and

heavy-atom root mean squared deviations for the structured segments in Fig. 2 are 1.2 and 1.8 Å, respectively. **c**, Chemical structure of the spin-labelled GDP, with the paramagnetic nitroxide moiety circled in red. **d**, Mapping the effect of spin-labelled GDP onto the ribbon drawing of UCP2. The colour gradient is from yellow (resonance intensity ratio of broadened to non-broadened,  $\epsilon = 1.0$ ) to white ( $\epsilon = 0.8$ ) to blue ( $\epsilon = 0.3$ ).



**Figure 4 | Comparison of UCP2 and ANT1.** **a**, Side views of UCP2 and ANT1 (Protein Data Bank ID, 1OKC). Pseudo-repeats one (residues 14–112), two (113–210) and three (211–309) are shown in blue, green and pink, respectively (see Fig. 3a for reference). **b**, Views of UCP2 and ANT1 from the matrix side of the carriers, showing loss of threefold pseudo-symmetry in UCP2 as a result of structural differences in repeat three. The orientations of the amphipathic helices are indicated by the arrows.

against the segment of the odd-numbered TMH that follows the conserved proline, and the proline kink in the TMH closes the channel (Fig. 4b). Repeat three of the GDP-bound UCP2 does not follow this pattern: TMH5 seems to have shifted between TMH4 and TMH6 towards the intermembrane side of the carrier (Supplementary Fig. 8). The APH of the same repeat rotates away from the channel by  $\sim 45^\circ$ , and its flanking regions are also substantially different from their counterparts in the other two repeats (Fig. 4b). Consequently, the matrix side of the channel is substantially more open in UCP2 than in ANT1. Although their functional relevance remains to be investigated, the differences we see could have implications for channel opening in proteins of the carrier family.

More than 40 different mitochondrial carriers have so far been identified; they transport a variety of metabolites, nucleotides, ions and vitamins across the inner mitochondrial membrane. Although their overall folds are expected to be similar, obtaining structures of different carriers and different functional states of the same carrier will be important for understanding the mechanism and selectivity of substrate transport<sup>29</sup>. Use of RDCs is well suited to this purpose—for example for describing concerted rearrangements of helices, coupled to opening or closing of the carrier on either side of the membrane—as RDCs are sensitive to the relative orientation of structural segments. We emphasize that new methods of measuring side-chain constraints must be developed to provide higher-resolution NMR structures of these carriers and of membrane proteins of comparable size.

## METHODS SUMMARY

Mouse UCP2 (residues 14–309, with a C-terminal His<sub>6</sub> tag) was expressed using a pET-21 vector in *Escherichia coli* Rosetta DE3 cells. After cell lysis, the lipid

composition of the membrane fraction was adjusted by adding DMPC, cardiolipin and phytanoyl lipids. The protein was extracted using 0.2% DPC in the presence of GDP. The solubilized UCP2 was then subjected to a series of purification steps including Ni-NTA affinity, MonoQ ion exchange, nucleotide-analogue affinity and size exclusion chromatography. The final NMR sample contained 0.8 mM UCP2, 5 mM GDP, 150 mM DPC, 2 mM DMPC, 1 mM cardiolipin, 5 mM  $\beta$ -mercaptoethanol, 30 mM potassium phosphate (pH 6.5), and 80 mM NaCl.

NMR experiments were conducted at 33 °C on spectrometers equipped with cryogenic probes. Sequence-specific assignment of backbone chemical shifts was accomplished using three pairs of triple-resonance experiments and double  $^{15}\text{N}$ -edited nuclear Overhauser enhancement spectroscopy, recorded on protein labelled with  $^{15}\text{N}$ ,  $^{13}\text{C}$  and  $^2\text{H}$ . RDCs were obtained using an aligned sample containing 0.5 mM UCP2 and 20 mg ml<sup>-1</sup> DNA nanotubes<sup>19</sup> (other components were the same as above).  $^1D_{\text{NH}}$  was measured by *J*-scaled TROSY-HNCO.  $^1D_{\text{C}'\text{C}\alpha}$  and  $^1D_{\text{NC}'}$  were measured by TROSY-HNCO with quantitative modulations of  $J_{\text{C}'\text{C}\alpha}$  (the coupling between backbone C' and C $\alpha$  spins) and  $J_{\text{NC}'}$  (the coupling between backbone N and C' spins), respectively. To obtain PREs, we generated a Cys-less UCP2 mutant and introduced single cysteines at desired positions for labelling with MTSL (Methods). Residue-specific broadening of protein resonances was measured with two TROSY-HNCO spectra, one recorded after nitroxide labelling and another after reducing the nitroxide free electron with ascorbic acid.

Structure determination had two stages: determination of local structural segments using an RDC-based MFR protocol and determination of the spatial arrangement of the MFR-derived segments using PRE distance restraints. Structures were calculated using XPLOR-NIH<sup>30</sup> with backbone  $\phi$  and  $\psi$  values of the assigned structured segments, RDCs and PRE-derived distances. A total of 30 structures were calculated using a simulated annealing protocol, and 15 low-energy structures were selected as the structural ensemble (statistics in Supplementary Table 1).

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 26 December 2010; accepted 1 June 2011.

Published online 24 July 2011.

- Palmieri, F. *et al.* Identification of mitochondrial carriers in *Saccharomyces cerevisiae* by transport assay of reconstituted recombinant proteins. *Biochim. Biophys. Acta* **1757**, 1249–1262 (2006).
- Krauss, S., Zhang, C. Y. & Lowell, B. B. The mitochondrial uncoupling-protein homologues. *Nature Rev. Mol. Cell Biol.* **6**, 248–261 (2005).
- Tate, C. G. & Stevens, R. C. Growth and excitement in membrane protein structural biology. *Curr. Opin. Struct. Biol.* **20**, 399–400 (2010).
- Peabay-Peyroula, E. *et al.* Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractylolide. *Nature* **426**, 39–44 (2003).
- Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144–148 (1961).
- Aquila, H., Link, T. A. & Klingenberg, M. The uncoupling protein from brown fat mitochondria is related to the mitochondrial ADP/ATP carrier. Analysis of sequence homologies and of folding of the protein in the membrane. *EMBO J.* **4**, 2369–2376 (1985).
- Zhang, C. Y. *et al.* Uncoupling protein-2 negatively regulates insulin secretion and is a major link between obesity, beta cell dysfunction, and type 2 diabetes. *Cell* **105**, 745–755 (2001).
- Diao, J. *et al.* UCP2 is highly expressed in pancreatic alpha-cells and influences secretion and survival. *Proc. Natl Acad. Sci. USA* **105**, 12057–12062 (2008).
- Andrews, Z. B. *et al.* UCP2 mediates ghrelin's action on NPY/AgRP neurons by lowering free radicals. *Nature* **454**, 846–851 (2008).
- Harper, M. E. *et al.* Characterization of a novel metabolic strategy used by drug-resistant tumor cells. *FASEB J.* **16**, 1550–1557 (2002).
- Samudio, I., Fiegl, M. & Andreeff, M. Mitochondrial uncoupling and the Warburg effect: molecular basis for the reprogramming of cancer cell metabolism. *Cancer Res.* **69**, 2163–2166 (2009).
- Schnell, J. R. & Chou, J. J. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **451**, 591–595 (2008).
- Hiller, S. *et al.* Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science* **321**, 1206–1210 (2008).
- Wang, J., Pielak, R. M., McClintock, M. A. & Chou, J. J. Solution structure and functional analysis of the influenza B proton channel. *Nature Struct. Mol. Biol.* **16**, 1267–1271 (2009).
- Zhou, Y. *et al.* NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. *Mol. Cell* **31**, 896–908 (2008).
- Van Horn, W. D. *et al.* Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science* **324**, 1726–1729 (2009).
- Gautier, A., Mott, H. R., Bostock, M. J., Kirkpatrick, J. P. & Nietlispach, D. Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nature Struct. Mol. Biol.* **17**, 768–774 (2010).
- Tjandra, N. & Bax, A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278**, 1111–1114 (1997).

19. Douglas, S. M., Chou, J. J. & Shih, W. M. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proc. Natl Acad. Sci. USA* **104**, 6644–6648 (2007).
20. Jones, T. A. & Thirup, S. Using known substructures in protein model building and crystallography. *EMBO J.* **5**, 819–822 (1986).
21. Delaglio, F., Kontaxis, G. & Bax, A. Protein structure determination using molecular fragment replacement and NMR dipolar couplings. *J. Am. Chem. Soc.* **122**, 2142–2143 (2000).
22. Shen, Y. *et al.* Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl Acad. Sci. USA* **105**, 4685–4690 (2008).
23. Raman, S. *et al.* NMR structure determination for larger proteins using backbone-only data. *Science* **327**, 1014–1018 (2010).
24. Losonczi, J. A., Andrec, M., Fischer, M. W. F. & Prestegard, J. H. Order matrix analysis of residual dipolar couplings using singular value decomposition. *J. Magn. Reson.* **138**, 334–342 (1999).
25. Cornilescu, G., Marquardt, J. L., Ottiger, M. & Bax, A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **120**, 6836–6837 (1998).
26. Battiste, J. L. & Wagner, G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear Overhauser effect data. *Biochemistry* **39**, 5355–5365 (2000).
27. Wang, Y. & Tajkhorshid, E. Electrostatic funneling of substrate in mitochondrial inner membrane carriers. *Proc. Natl Acad. Sci. USA* **105**, 9598–9603 (2008).
28. Dehez, F., Pebay-Peyroula, E. & Chipot, C. Binding of ADP in the mitochondrial ADP/ATP carrier is driven by an electrostatic funnel. *J. Am. Chem. Soc.* **130**, 12725–12733 (2008).
29. Kunji, E. R. & Robinson, A. J. Coupling of proton and substrate translocation in the transport cycle of mitochondrial carriers. *Curr. Opin. Struct. Biol.* **20**, 440–447 (2010).
30. Schwieters, C. D., Kuszewski, J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 66–74 (2002).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank K. Oxenoid and R. Sounier for discussions, M. McClintock for help with DNA nanotube preparation, I. Stokes-Rees and P. Sliz for help with computations, and N. Voigt for help with figures. The work was supported by NIH grants 1U54GM094608 (to J.J.C.) and 1DP2OD004641 (to W.M.S.). S.C.H. is an Investigator at the Howard Hughes Medical Institute.

**Author Contributions** M.J.B., W.M.S., S.C.H. and J.J.C. designed the study; M.J.B. prepared NMR samples; M.J.B. and W.M.S. prepared DNA nanotubes; M.J.B. and J.J.C. designed experiments, collected and analysed NMR data, and determined the structure; M.J.B. and J.J.C. wrote the paper; and all authors contributed to editing the manuscript.

**Author Information** The structure of UCP2 has been deposited in the Protein Data Bank under accession number 2LCK. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to J.J.C. ([chou@cmcd.hms.harvard.edu](mailto:chou@cmcd.hms.harvard.edu)).



## METHODS

**Sample preparation.** Mouse UCP2 (residues 14–309 with a C-terminal His<sub>6</sub> tag) was expressed using a pET-21 vector in *E. coli* Rosetta DE3 cells. After cell lysis, Triton X-100 was added to the lysate at 10  $\mu$ l per gram of cell pellet, followed by stirring at 4 °C for 2 h in the presence of 10  $\mu$ M GDP and 0.2% BME. The suspension was homogenized and centrifuged at 1,500g to remove insoluble debris. To adjust the lipid composition of the membrane fraction, 10 mg DMPC, 2 mg cardiolipin and 0.1 mg phytanoyl lipid per gram of cell pellet were solubilized at 20 mg ml<sup>-1</sup> in 10% octyl glucoside and then added to the supernatant, followed by 2 h of stirring at 4 °C. The membranes were collected by centrifugation at 50,000g for 2 h, and UCP2 was extracted in 40 mM potassium phosphate (pH 8.0), 250 mM NaCl, 50 mM BME, 10  $\mu$ M GDP and 0.2% DPC. The solubilized UCP2 was adsorbed on a Ni-NTA resin in the absence of GDP and eluted with 300 mM imidazole. The sample was then exchanged by dialysis into low-salt buffer and applied to a MonoQ column. The MonoQ flowthrough was incubated with Ciba Blue nucleotide-analogue resin (Sigma) for 1 h, eluted with 100  $\mu$ M GDP and, finally, loaded onto a Sephadex S-200 column for size exclusion purification in 50 mM potassium phosphate (pH 6.5), 100 mM NaCl and 5 mM DPC. The eluted UCP2 sample was supplemented with GDP, detergent and lipids such that the final NMR sample contained 0.8 mM UCP2, 5 mM GDP, 150 mM DPC, 2 mM DMPC, 1 mM cardiolipin, 5 mM BME, 30 mM potassium phosphate (pH 6.5) and 80 mM NaCl.

We have also tested expression and purification of mouse UCP1 and UCP3. UCP1 could be expressed and purified as described for UCP2, and showed FFA-dependent proton translocation activity and GDP inhibition. UCP3 could be expressed but presented serious solubility problems during purification.

**GDP binding.** We used FRET to test GDP binding to UCP2 under the NMR conditions. The intrinsic tryptophan fluorescence ( $\lambda_{\text{ex}} = 280$  nm,  $\lambda_{\text{em}} = 350$  nm) of UCP2 was the resonance energy transfer donor and MANT-GDP ( $\lambda_{\text{ex}} = 356$  nm,  $\lambda_{\text{em}} = 440$  nm) was the acceptor. We recorded the emission intensity at 440 nm ( $\lambda_{\text{ex}} = 280$  nm) of 10  $\mu$ M UCP2/MANT-GDP complex for increasing GDP concentrations. Figure 1c shows the FRET response as a percentage relative to 1:0 MANT-GDP:GDP. Data were fitted to the standard equilibrium binding equation.

**NMR spectroscopy.** NMR experiments were conducted at 33 °C on spectrometers equipped with cryogenic TXI probes (Bruker). Sequence-specific assignment of backbone <sup>1</sup>H<sup>N</sup>, <sup>15</sup>N, <sup>13</sup>C $\alpha$ , <sup>13</sup>C $\beta$ , and <sup>13</sup>C' chemical shifts were accomplished using three pairs of triple-resonance experiments and a double <sup>15</sup>N-edited NOESY, recorded on a protein sample labelled with <sup>15</sup>N, <sup>13</sup>C and <sup>2</sup>H. The triple-resonance experiments include the TROSY versions of HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HN(CA)CO and HNCO<sup>31,32</sup>. The <sup>15</sup>N-edited NOESY is a high-resolution, 3D (H<sup>N</sup>, H<sup>N</sup>)-HMQC-NOESY-TROSY with <sup>15</sup>N, <sup>15</sup>N and <sup>1</sup>H<sup>N</sup> evolution in the *t*<sub>1</sub>, *t*<sub>2</sub> and *t*<sub>3</sub> dimensions, respectively.

The RDCs were obtained by subtracting *J* of the isotropic sample above from *J* + *D* of the aligned sample containing 0.5 mM protein and 20 mg ml<sup>-1</sup> DNA nanotubes (all other components were the same as in the isotropic sample). Details of preparing DNA nanotubes and aligned protein samples are in Supplementary Information. The sign of the RDC follows the convention that  $|J_{\text{NH}} + {}^1D_{\text{NH}}| < 92$  Hz when  ${}^1D_{\text{NH}}$  is positive. The <sup>1</sup>H–<sup>15</sup>N couplings were measured at 600 MHz (<sup>1</sup>H frequency) using the *J*-scaled TROSY-HNCO experiment to exploit the favourable relaxation property of the TROSY transitions<sup>33</sup>. In this experiment, two interleaved spectra were recorded, the regular TROSY-HNCO spectrum and a modified TROSY-HNCO spectrum with *J*<sub>NH</sub> evolution during the <sup>15</sup>N chemical shift evolution scaled to zero. The <sup>13</sup>C'–<sup>13</sup>C $\alpha$  couplings were measured at 600 MHz using the 3D TROSY-HNCO experiment with quantitative *J*<sub>C'C $\alpha$</sub>  modulations of 0 and 28 ms (ref. 34). The <sup>13</sup>C'–<sup>15</sup>N couplings were measured at 750 MHz using the 3D TROSY-HNCO experiment with quantitative *J*<sub>NH</sub> modulations of 33 and 66 ms (ref. 35).

To introduce a single paramagnetic site for PRE measurement, the five cysteines of UCP2 were all mutated to alanine or serine (Cys 25 Ala, Cys 191 Ser, Cys 217 Ser, Cys 227 Ser, Cys 256 Ser); the cysteine-free UCP2 had GDP binding properties similar to those of wild-type protein. A single cysteine was introduced into the protein on the basis of the known secondary structures from MFR segments. The

nitroxide bearing compound, *S*-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL), was then attached at the cysteine position by adding sevenfold excess label to 10  $\mu$ M UCP2 in the NMR buffer at pH 8.0, and incubating at 25 °C for 4 h. Excess label was removed to avoid nonspecific broadening. The pH was changed back to 6.5 for NMR measurements. To quantify residue-specific broadening of backbone <sup>1</sup>H<sup>N</sup>, two TROSY-HNCO spectra were recorded, one after nitroxide labelling and the other after reduction of the nitroxide free electron with a  $\times 5$  molar ascorbic acid. The same protocol was used to measure broadening of UCP2 resonances by the spin-labelled GDP (Fig. 3c; custom synthesized by Shanghai ChemPartner).

**Structure determination.** The procedure had two stages: determination of local structural segments by RDC-based MFR and determination of the spatial arrangement of the MFR-derived segments in the tertiary fold using PRE distance restraints. In stage one, we fitted RDCs of every seven-residue stretch along the UCP2 sequence to the fragments in the database using the program PALES<sup>36</sup>. Only those seven-residue stretches with on average more than two RDCs per residue were used for fragment searching. In the end, 5,520 fragments with *Q*<sub>free</sub> < 0.25 were collected. For each candidate fragment (residue *n* – *m*) evaluated in the fragment assignment, gap filling or end extension operation (described in the main text), only the  $\phi$  and  $\psi$  angles of residue (*n* + 1) – (*m* – 1) were used, because the end residues do not have the <sup>15</sup>N–<sup>13</sup>C' dipolar coupling. During fragment assignment, we found that for each seven-residue stretch, the best 5–10 fragments typically have very similar *Q*<sub>free</sub> values (differences less than 10%). Therefore, the best fragment was assigned to the region only if these fragments have a backbone root mean squared deviation of less than <0.5 Å. Similarly, there are multiple fragments that score equally well in gap filling or end extension. In those cases, we assigned the best-scoring fragment only if the final merged or extended segments have backbone root mean squared deviation of less than <1 Å. Examples of fragment convergence are shown in Supplementary Fig. 9.

In stage two, we calculated structures using XPLOR-NIH<sup>37</sup> with the assigned structured segments, RDCs and PREs. The  $\phi$  and  $\psi$  values of the segments in Fig. 2c were strongly enforced by a harmonic potential with force constant ramped from 10 to 1,000 kcal mol<sup>-1</sup> rad<sup>-2</sup>. All RDCs used for determining the segments were applied, and the RDC force constant was ramped from 0.01 to 1.5 kcal mol<sup>-1</sup> Hz<sup>-2</sup>. PRE restraints (uncertainties in Supplementary Table 2) were enforced with flat-well harmonic potentials, with the force constant ramped from 1 to 40 kcal mol<sup>-1</sup> Å<sup>-2</sup>. In addition to experimental restraints, a weak database-derived 'Rama' potential function<sup>38</sup> was ramped from 0.02 to 0.2 (dimensionless force constant) for the general treatment of side-chain rotamers. A total of 30 monomer structures were calculated using a simulated annealing protocol in which the bath temperature was cooled from 2,000 to 200 K. Fifteen low-energy structures were selected as the structural ensemble (Supplementary Table 1).

- Salzmann, M., Wider, G., Pervushin, K. & Wuthrich, K. Improved sensitivity and coherence selection for [N-15,H-1]-TROSY elements in triple resonance experiments. *J. Biomol. NMR* **15**, 181–184 (1999).
- Kay, L. E., Torchia, D. A. & Bax, A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 (1989).
- Kontaxis, G., Clore, G. & Bax, A. Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *J. Magn. Reson.* **143**, 184–196 (2000).
- Jaroniec, C. P., Ulmer, T. S. & Bax, A. Quantitative J correlation methods for the accurate measurement of <sup>13</sup>C'–<sup>13</sup>C $\alpha$  dipolar couplings in proteins. *J. Biomol. NMR* **30**, 181–194 (2004).
- Chou, J. J., Delaglio, F. & Bax, A. Measurement of one-bond <sup>15</sup>N–<sup>13</sup>C' dipolar couplings in medium sized proteins. *J. Biomol. NMR* **18**, 101–105 (2000).
- Zweckstetter, M. & Bax, A. Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. *J. Am. Chem. Soc.* **122**, 3791–3792 (2000).
- Schwieters, C. D., Kuszewski, J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 66–74 (2002).
- Kuszewski, J., Gronenborn, A. M. & Clore, G. M. Improvements and extensions in the conformational database potential for the refinement of NMR and X-ray structures of proteins and nucleic acids. *J. Magn. Reson.* **125**, 171–177 (1997).

# Transient dynamics of an altered large marine ecosystem

Kenneth T. Frank<sup>1</sup>, Brian Petrie<sup>1</sup>, Jonathan A. D. Fisher<sup>2</sup> & William C. Leggett<sup>2</sup>

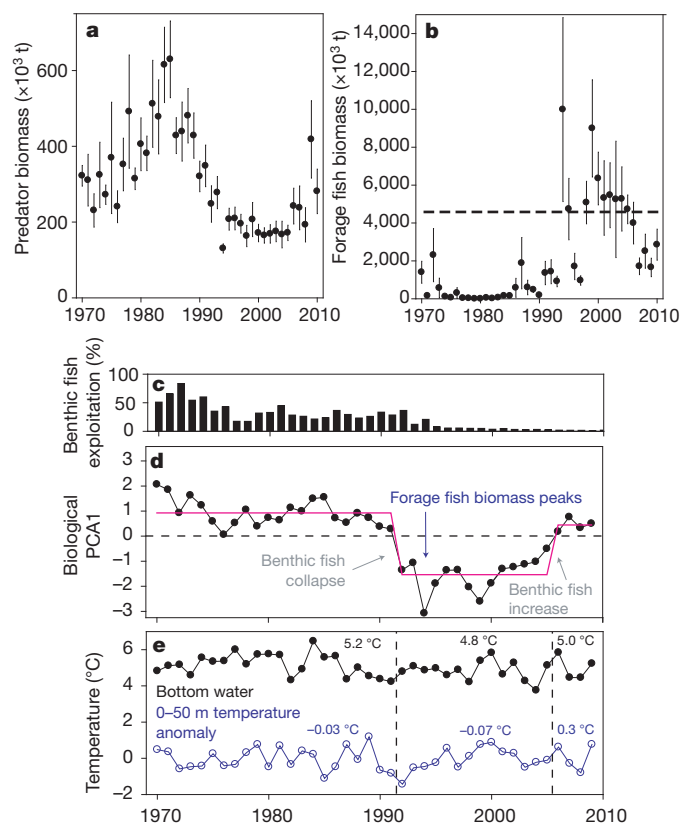
**Overfishing of large-bodied benthic fishes and their subsequent population collapses on the Scotian Shelf of Canada's east coast<sup>1,2</sup> and elsewhere<sup>3,4</sup> resulted in restructuring of entire food webs now dominated by planktivorous, forage fish species and macroinvertebrates. Despite the imposition of strict management measures in force since the early 1990s, the Scotian Shelf ecosystem has not reverted back to its former structure. Here we provide evidence of the transient nature of this ecosystem and its current return path towards benthic fish species domination. The prolonged duration of the altered food web, and its current recovery, was and is being governed by the oscillatory, runaway consumption dynamics of the forage fish complex. These erupting forage species, which reached biomass levels 900% greater than those prevalent during the pre-collapse years of large benthic predators, are now in decline, having outstripped their zooplankton food supply. This dampening, and the associated reduction in the intensity of predation, was accompanied by lagged increases in species abundances at both lower and higher trophic levels, first witnessed in zooplankton and then in large-bodied predators, all consistent with a return towards the earlier ecosystem structure. We conclude that the reversibility of perturbed ecosystems can occur and that this bodes well for other collapsed fisheries.**

The recent demonstration that overfishing of large-bodied predators in the northwest Atlantic initiated a trophic cascade, typified by reciprocal changes in biomass between adjacent trophic levels extending to the base of the food web<sup>1,2</sup>, overturned the long-held view that large marine ecosystems are resistant to restructuring<sup>5</sup>. It has been proposed<sup>6,7</sup> that such trophic cascades are characteristic of ecosystems that have been transformed into undesirable states involving large changes in ecological functions and/or economic resources<sup>8,9</sup>. Although the excessive consumption characteristic of trophic cascades may be unstable<sup>6</sup>, whether, how, and on what time scales such altered, diverse food webs and their key species and functional groups will recover remains unknown<sup>1,3,4</sup>. This has led to controversy regarding the efficacy of and experimentation with strategies based on conventional management approaches such as moratoria on exploitation, culling and re-stocking intended to return ecosystems to their former structure<sup>10–13</sup>. Using four decades of high quality, annual, fishery-independent data (see Methods) representative of multiple trophic levels on the eastern Scotian Shelf (Supplementary Fig. 1), we document the transient nature of its altered ecosystem and its return towards dominance by large-bodied predators.

The collapse of the northwest Atlantic cod (*Gadus morhua*) and several other large predatory fishes in the early 1990s (Fig. 1a), caused principally by over-fishing<sup>14,15</sup>, precipitated the first documented open ocean trophic cascade in a large marine ecosystem<sup>1</sup>. The total biomass of cod, one of the ecosystem's dominant species, has hovered at less than 5% of pre-collapse levels for almost two decades despite the implementation of strict regulations forbidding their capture<sup>15</sup>.

Recent investigations<sup>13,16</sup> have provided strong evidence that, following these collapses, the eastern Scotian Shelf, and other northwest Atlantic ecosystems in which similar collapses occurred, moved to

apparent alternate states in which planktivorous forage fishes and macroinvertebrates became the dominant predators<sup>3</sup>. Released from predation on the eastern Scotian Shelf, the biomass of forage fishes increased by 900% (Fig. 1b) and macroinvertebrates by 200% compared to the pre-collapse years<sup>13</sup>. They then competed directly with and/or preyed upon the early life stages of their once benthic predators, a phenomenon termed predator–prey reversal<sup>17</sup> which seems to be one of the leading causes of the delayed recovery of the benthic fish complex in this and other large marine ecosystems<sup>4,17,18</sup>. Although forage fish constitute approximately half the diet of an expanding, resident grey seal (*Halichoerus grypus*) population, estimates of their consumption of pelagic fish species (1995–2000) were only 35% of the benthic



**Figure 1 | Variability of the eastern Scotian Shelf ecosystem.** a–e, Data (Supplementary Fig. 1) based on large-bodied benthic fish (a), their prey (forage fishes, b) with estimated carrying capacity (dashed line), benthic fish exploitation history expressed as annual per cent biomass removal (c), changing ecosystem structure based on the leading mode (PCA1) of biotic data spanning four trophic levels and the demarcation of regimes<sup>29</sup> of 22, 14 and 4 years duration (pink solid line) (d), and temperatures with averages shown for the three regimes (dashed vertical lines) (e). Vertical bars in a and b show  $\pm$  s.e.m. ( $n = 27$ ).

<sup>1</sup>Ocean Sciences Division, Bedford Institute of Oceanography, P.O. Box 1006, Dartmouth, Nova Scotia B2Y 4A2, Canada. <sup>2</sup>Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada.

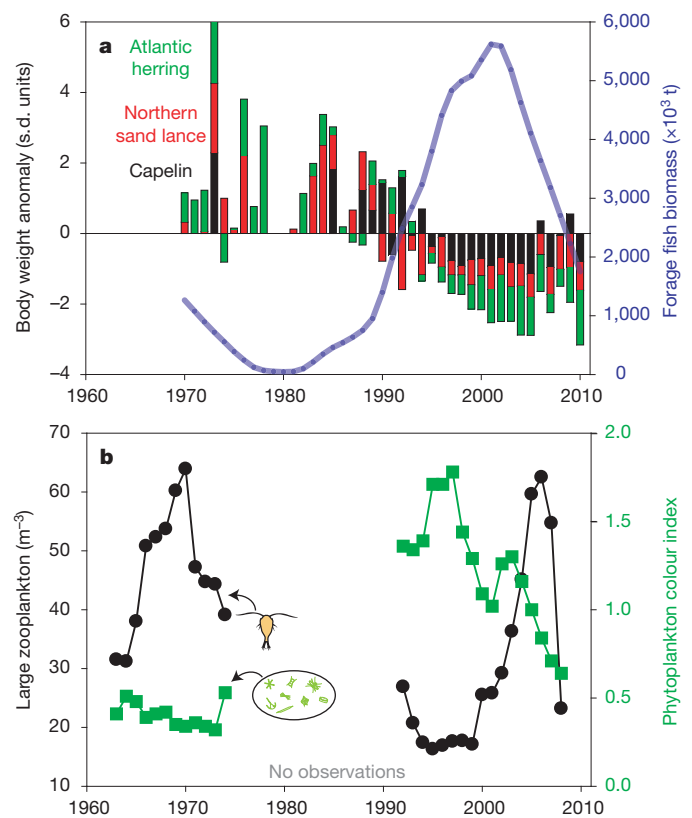
fish complex<sup>13</sup> and insufficient to suppress the outbreaks and biomass variability of forage fishes (Fig. 1b).

The changing status of the eastern Scotian Shelf ecosystem exhibited two transitions. A period of intensive fishing, when aggregate landings of benthic fishes averaged close to 105,000 tonnes (t) representing annual biomass removals of >50% (Fig. 1c), resulted in the first transition centred in 1991–1992 from a ‘pre-collapse’ state dominated by benthic fish species to a ‘collapsed’ state dominated by forage fish species<sup>1</sup>. A cod and haddock fishing moratorium implemented in 1993 had the desired effect of reducing aggregate exploitation (<5% since 2000; Fig. 1c) but did not produce the anticipated recovery (Supplementary Figs 2 and 3). The second transition, centred in 2005–2006, represents a return towards a ‘recovering’ state of benthic fish domination described here (Fig. 1d and Methods). An additional bottom trawl survey, beginning in 1986, revealed a similar pattern of collapse and recent increase in benthic fish biomass (Supplementary Fig. 4).

The physical environment during the three states, assessed using annual bottom temperatures, 0–50 m water temperatures (Fig. 1e) and water column stratification (Supplementary Information), showed only minor changes. The bottom temperatures during the collapsed state were only 0.33 °C and 0.24 °C lower than during the pre-collapse and recovering periods, respectively. The magnitude of this temperature change would have only minimal or no effect on individual and population growth rates as well as other life history traits (Methods). Further, the dominant large-scale atmospheric forcing mechanism in the western North Atlantic (that is, the North Atlantic Oscillation) induces a bimodal response of ocean temperatures with a nodal point in bottom temperature occurring in the middle of the eastern Scotian Shelf<sup>9</sup>. Consequently, the temperature response to such forcing in this region is minimal and is reflected in the dampening of regional variability in other biological properties such as species richness<sup>20</sup>. Differences in water column temperature anomalies were also slight for the three periods: on average, temperatures during the pre-collapse and collapsed periods were within 0.1 °C, during the forage fish outbreak temperatures were elevated by about 0.4 °C; overall, temperatures varied over a range of about 2 °C. There was no relationship between water column temperatures and forage fish biomass at zero lag (Supplementary Fig. 5; correlation coefficient,  $r = 0.02$ ) or for lags (forage fish biomass relative to temperature) up to four years (Methods). The minor increases in stratification occurred primarily during summer, outside of the peak period of phytoplankton production (Methods). The timing and magnitude of this ongoing recovery of the benthic fish complex was initiated and is being sustained by naturally induced changes in the dynamics of their former prey, and the resulting impact on the total ecosystem, more so than by external climatic influences (Fig. 1e).

The second and most recent ecosystem transition began with a marked decline in the biomass of the unfished, forage fish complex dominated by northern sand lance (*Ammodytes dubius*), capelin (*Mallotus villosus*), and Atlantic herring (*Clupea harengus*). The aggregate biomass of these species peaked in 1994 and 1999 at approximately 10 million t, which exceeded the estimated carrying capacity of 4.3 million t for the eastern Scotian Shelf ecosystem (Fig. 1b; Methods). Subsequently, their total biomass rapidly declined at an average rate of 0.5 million t per year to current levels near 3 million t. Such eruptions followed by crashes involving fast growing, highly opportunistic species are known to occur in other ecosystems freed from predatory control<sup>21,22</sup>.

Physiological changes and cascading food web effects associated with the overshoot of the pelagic forage fish were evident. Relative weights, an index of physiological condition available since 1970, of the three dominant forage fish species showed coherent changes (31 out of 41 years with same sign) throughout the entire period with relative high condition indices from the 1970s to the early 1990s followed by sustained declines beginning about 1994 (Fig. 2a). This points to the pelagic species having inadequate food resources at

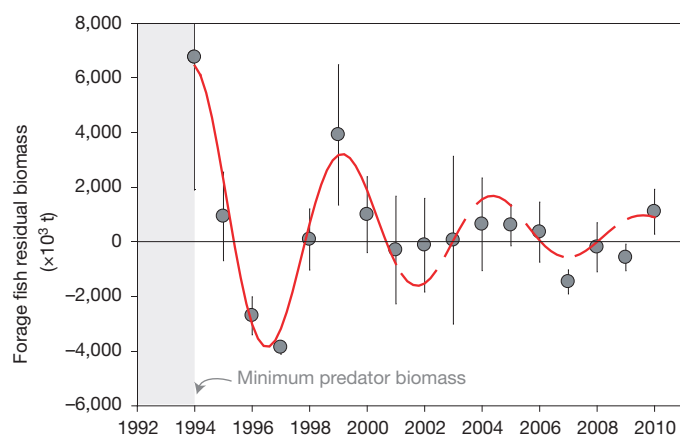


**Figure 2 | Physiological changes in forage fish species and resultant food web effects.** **a**, Species-specific body weight anomalies (stacked histograms) and smoothed forage fish biomass (solid line; 30% LOESS; Supplementary Fig. 8). **b**, Time series of large-bodied zooplankton abundance and phytoplankton colour index (a measure of abundance) from the Continuous Plankton Recorder survey (<http://www.safhos.ac.uk>). Recent (2000–2007) increases in large-bodied zooplankton, coincident with major reductions in forage fish biomass shown in **a**, indicate a weakening of the inhibitory effects of the predator–prey reversal mechanism to benthic fish species recovery.

increased abundances. The density of large-bodied zooplankton, which has varied inversely ( $r = -0.32$ ,  $n = 22$  years) with the forage fish biomass (their principal predators), reached a broad minimum of about 17 individuals per  $m^3$  at the approximate peak (1994) of the forage group biomass—a signature of excessive grazing. Large-bodied zooplankton species increased rapidly (1997–2007) by a factor of four in six to seven years to levels observed before the forage fish eruption, the final year in the series being an exception (Fig. 2b). Moreover, the standing stock of phytoplankton, which has varied inversely with the abundance of large-bodied zooplankton ( $r = -0.71$ ), declined by 40% during the ongoing recovery of the benthic fish complex. These reciprocal relationships between adjacent trophic levels are consistent with the trophic cascade model<sup>1</sup>.

Most revealing is the fact that the trajectory of forage fish biomass changes following the benthic fish collapse approximates a damped harmonic oscillator (Fig. 3), providing key biological insights from the derived parameter values including estimates of the period, 5 years (the approximate life cycle of this functional group), and dampening time, 7 years. The 7-year dampening represented a 78% decrease in forage fish biomass between 1994, the time of the peak amplitude, and 2005, the end of the regime dominated by this group (Fig. 1d). Studies of terrestrial herbivores indicate that the cycle period is strongly dependent on body size<sup>22</sup>; based on the range of body sizes of the forage fishes we studied (weight range: 0.04–0.3 kg), the cycle period is similar to those estimated from outbreaks of similar-sized terrestrial species (3.6 to 5.9 years)<sup>22</sup>. This dynamic, oscillatory behaviour indicates that the internal damping capacity of the ecosystem, and not solely management





**Figure 3 | Post-perturbation forage fish variability.** The variability (solid points  $\pm$  s.e.m.,  $n = 27$ ), described by a damped harmonic oscillator fitting the 1994–2001 observations (solid line), gives an amplitude of  $6,100 \times 10^3$  t, dampening rate = 6.9 years, and oscillation period = 5.2 years (Methods). Shaded area indicates period of greatest population growth. Aggregate prey biomass is expressed as an anomaly from the 30% LOESS filter (Supplementary Fig. 8). This solution is extrapolated to the end of the record (broken line). In theory<sup>30</sup>, a population or functional group that is characterized by damped oscillations will overshoot and then undershoot the carrying capacity.

strategies, probably initiated the return of the eastern Scotian Shelf ecosystem towards its former structure and the restoration of food web stability.

We propose that this crash in forage fish biomass led to a reduction in the intensity of predator–prey reversal which, combined with increased food availability to the benthic fish larval stages resulting from the corresponding and related growth in the abundance of large zooplankton, provided a window of opportunity for the recovery of the once dominant benthic predators. Since 2006 the benthic fish biomass has attained levels approaching those observed during the pre-collapse period (Fig. 1a). Atlantic cod and redfish (*Sebastes* spp.) have reached levels not seen since the early 1990s and haddock (*Melanogrammus aeglefinus*) to an unprecedented high (Supplementary Fig. 6). Enhanced recruitment (by a factor of 5.3, recovering/pre-collapse; Supplementary Fig. 7) and improvements in post-recruitment survivorship, which for cod and haddock has increased by factors of 12 and 70 times compared to the collapsed period, contributed to these changes (Supplementary Table 1). In addition, three of the four benthic fish species which are routinely aged (cod, pollock and silver hake) have shown 8–18% increases in average weight at age during 2006–2010 relative to the 1992–2005 post-collapse period (Supplementary Table 2).

The generally positive response of the large-bodied zooplankton to the declining forage fish biomass supports this hypothesis. The early life stages of most of the benthic fish species form an integral part of the large-bodied zooplankton complex and their survival and contribution to recruitment would have benefited from the same forces that have led to the increase in large-bodied zooplankton abundance noted above. Increased predation by the expanding benthic fish complex on the forage fish community should accelerate this trend.

This unfolding drama held many surprises, including the prolonged recovery of the benthic predator complex, despite the moratorium on directed fishing for cod and haddock, the establishment of a closed area on the western offshore banks that preceded the fishing moratorium<sup>23</sup>, and the promotion of new and experimental fisheries<sup>24</sup> to divert fishing effort away from the traditional species. Although the current trajectory is positive, several factors could alter ongoing ecosystem recovery. The current high levels of recruitment and survivorship of the benthic fish complex, if sustained, could accelerate the recovery. The current dominance of haddock over cod also raises the question of whether the species makeup of the ecosystem will return to that which prevailed before the collapse. Furthermore, recovery in other over-exploited

ecosystems such as the Black Sea, Northern Benguela, the Sea of Japan, and elsewhere has been delayed by jellyfish blooms<sup>25</sup>, the presence of invasive species and by eutrophication<sup>3</sup>, all of which are possible in the system we describe. The widespread body size reductions of benthic fishes documented for other exploited northwest Atlantic systems<sup>26,27</sup>, if not reversed, could also slow the recovery of the benthic fish complex and adversely affect food web structure<sup>26</sup>. The evolving global climate could alter the ecosystem positively or negatively.

These uncertainties notwithstanding, the answer to the critical question of whether or not such profound changes in the dynamics of large marine ecosystems are reversible seems to be ‘yes’. This bodes well for other perturbed, formerly cod-dominated systems at latitudes to the north of the eastern Scotian Shelf that have yet to recover. Indeed, subtle signs of cod recovery have been appearing in other sub-arctic northwest Atlantic ecosystems during the past few years<sup>28</sup>. However, the time scales for their recovery are likely to be greater given the lower water temperatures (equates to slower turnover times) and their reduced species richness and, for some, because of the continued exploitation of cod and other large-bodied benthic fish species<sup>13,15,18</sup>.

## METHODS SUMMARY

Annual fishery-independent, randomly stratified bottom trawl surveys on the continental shelf off eastern Nova Scotia (1970–2010) during July–August provided biomass and variance estimates for functional groupings of fifteen, commercially exploited benthic and three forage fish species. A March survey (1986–2010) was used to assess the benthic fish dynamics further. Benthic fish exploitation levels were expressed as the ratio of landings to survey-estimated biomass. Ageing data, available for four benthic fish species, were used to calculate growth and mortality rates. Correction factors were applied to the typically under-sampled forage species. A 30% linear LOESS filter was applied to the forage fish biomass data to isolate high-frequency variability; the resultant anomalies were least squares fit to a damped harmonic oscillator equation (Methods). We quantified the time-averaged ecosystem carrying capacity of forage fishes using seasonally averaged, zooplankton data and production to biomass ratios obtained from the literature (Methods). Annual anomalies of the individual average body weights (total biomass/total abundance), an index of physiological condition, were derived for each forage species. Lower trophic level data (phytoplankton colour index and zooplankton abundance time series) were obtained from the Continuous Plankton Recorder survey (<http://www.safhos.ac.uk>). Zooplankton were grouped into large ( $\geq 2$  mm carapace length) and small ( $< 2$  mm) species. The eastern Scotian Shelf ecosystem status was assessed by principal components analysis of the five biological time series. To determine whether and when ecosystem transitions occurred, the dominant principal component axis (PCA1) was subjected to a sequential *t*-test analysis of the regime shift method (STARS)<sup>29</sup>, which identifies the magnitude and direction of significant shifts. Temperature, salinity and density observations were obtained from directed and opportunistic ship-based sampling.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 10 February; accepted 13 June 2011.**

**Published online 27 July 2011.**

1. Frank, K. T., Petrie, B., Choi, J. S. & Leggett, W. C. Trophic cascades in a formerly cod-dominated ecosystem. *Science* **308**, 1621–1623 (2005).
2. Frank, K. T., Petrie, B., Shackell, N. L. & Choi, J. S. Reconciling differences in trophic control in mid-latitude marine ecosystems. *Ecol. Lett.* **9**, 1096–1105 (2006).
3. Daskalov, G. M., Grishin, A. N., Rodionov, S. & Mihneva, V. Trophic cascades triggered by overfishing reveal possible mechanisms of ecosystem regime shifts. *Proc. Natl Acad. Sci. USA* **104**, 10518–10523 (2007).
4. Casini, M. et al. Multi-level trophic cascades in a heavily exploited open marine ecosystem. *Proc. R. Soc. Lond. B* **275**, 1793–1801 (2008).
5. Steele, J. H. & Collie, J. S. in *The Global Coastal Ocean: Multiscale Interdisciplinary Processes* (eds Robinson, A. R. & Brink, K.) vol. 13, chap. 21 (Harvard, 2004).
6. Strong, D. R. Are trophic cascades all wet? Differentiation and donor-control in species ecosystems. *Ecology* **73**, 747–754 (1992).
7. Pace, M. L., Cole, J. J., Carpenter, S. R. & Kitchell, J. F. Trophic cascades revealed in diverse ecosystems. *Trends Ecol. Evol.* **14**, 483–488 (1999).
8. Jackson, J. B. C. et al. Historical overfishing and the recent collapse of coastal ecosystems. *Science* **293**, 629–638 (2001).
9. Myers, R. A., Baum, J. K., Shepherd, T. D., Powers, S. P. & Peterson, C. H. Cascading effects of the loss of apex predatory sharks from a coastal ocean. *Science* **315**, 1846–1850 (2007).
10. Persson, L. et al. Culling prey promotes predator recovery—alternative states in a whole-lake experiment. *Science* **316**, 1743–1746 (2007).

11. Terborgh, J. & Estes, J. A. *Trophic Cascades: Predators, Prey, and the Changing Dynamics of Nature* (Island, 2010).
12. Yodzis, P. Must top predators be culled for the sake of fisheries? *Trends Ecol. Evol.* **16**, 78–84 (2001).
13. Bundy, A., Heymans, J. J., Morissette, L. & Savenkoff, C. Seals, cod and forage fish: a comparative exploration of variations in the theme of stock collapse and ecosystem change in four Northwest Atlantic ecosystems. *Prog. Oceanogr.* **81**, 188–206 (2009).
14. Myers, R. A. & Worm, B. Extinction, survival or recovery of large predatory fishes. *Phil. Trans. R. Soc. B* **360**, 13–20 (2005).
15. Shelton, P. A. *et al.* Fishing under low productivity conditions is further delaying recovery of Northwest Atlantic cod (*Gadus morhua*). *Can. J. Fish. Aquat. Sci.* **63**, 235–238 (2006).
16. Savenkoff, C. *et al.* Effects of fishing and predation in a heavily exploited ecosystem: comparing periods before and after the collapse of groundfish in the southern Gulf of St. Lawrence (Canada). *Ecol. Modell.* **204**, 115–128 (2007).
17. Fauchald, P. Predator–prey reversal: a possible mechanism for ecosystem hysteresis in the North Sea? *Ecology* **91**, 2191–2197 (2010).
18. Petrie, B., Frank, K. T., Shackell, N. L. & Leggett, W. C. Structure and stability in exploited marine fish communities: quantifying critical transitions. *Fish. Oceanogr.* **18**, 83–101 (2009).
19. Petrie, B. Does the North Atlantic Oscillation affect hydrographic properties on the Canadian Atlantic continental shelf? *Atm.–Ocean* **45**, 141–151 (2007).
20. Fisher, J. A. D., Frank, K. T., Petrie, B., Leggett, W. C. & Shackell, N. L. Temporal dynamics within a contemporary latitudinal diversity gradient. *Ecol. Lett.* **11**, 883–897 (2008).
21. Forsyth, D. M. & Caley, P. Testing the irruptive paradigm of large-herbivore dynamics. *Ecology* **87**, 297–303 (2006).
22. Peterson, R. O., Page, R. E. & Dodge, K. M. Wolves, moose, and the allometry of population cycles. *Science* **224**, 1350–1352 (1984).
23. Frank, K. T., Shackell, N. L. & Simon, J. E. An evaluation of the Emerald/Western Bank juvenile haddock closed area. *ICES J. Mar. Sci.* **57**, 1023–1034 (2000).
24. Anderson, S. C., Lotze, H. K. & Shackell, N. L. Evaluating the knowledge base for expanding low-trophic-level fisheries in Atlantic Canada. *Can. J. Fish. Aquat. Sci.* **65**, 2553–2571 (2008).
25. Richardson, A. J., Bakun, A., Hays, G. C. & Gibbons, M. J. The jellyfish joyride: causes, consequences and management responses to a more gelatinous future. *Trends Ecol. Evol.* **24**, 312–322 (2009).
26. Shackell, N. L., Frank, K. T., Fisher, J. A. D., Petrie, B. & Leggett, W. C. Decline in top predator body size and changing climate alter trophic structure in an oceanic ecosystem. *Proc. R. Soc. B* **277**, 1353–1360 (2010).
27. Fisher, J. A. D., Frank, K. T. & Leggett, W. C. Breaking Bergmann's rule: truncation of Northwest Atlantic marine fish body sizes. *Ecology* **91**, 2499–2505 (2010).
28. Murawski, S. A. Rebuilding depleted fish stocks: the good, the bad, and mostly, the ugly. *ICES J. Mar. Sci.* **67**, 1830–1840 (2010).
29. Rodionov, S. N. A sequential algorithm for testing climate regime shifts. *Geophys. Res. Lett.* **31**, L09204 (2004).
30. Rooney, N., McCann, K., Gellner, G. & Moore, J. C. Structural asymmetry and the stability of diverse food webs. *Nature* **442**, 265–269 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the Department of Fisheries and Oceans staff who collected and maintained the multi-trophic level and biogeochemical monitoring data, J. Black for assistance with computing and database issues, and A. Bundy and D. R. Strong for helpful criticisms. Symbols for Fig. 2b are courtesy of the Integration and Application Network (<http://www.ian.umces.edu/symbols>). This research was supported by Fisheries and Oceans Canada and Discovery grants from the Natural Sciences and Engineering Research Council of Canada (to K.T.F. and W.C.L.).

**Author Contributions** K.T.F. developed the overall concept of paper reported here. B.P. and K.T.F. developed the initial hypothesis and assembled the relevant data. Most analyses were performed by B.P. with assistance from J.A.D.F. and K.T.F. All authors contributed to the interpretations and synthesis of the analytical results.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to K.T.F. ([kenneth.frank@dfo-mpo.gc.ca](mailto:kenneth.frank@dfo-mpo.gc.ca)).

## METHODS

**Large-bodied benthic predators, aggregate forage fish biomass, zooplankton functional groups and chlorophyll.** The primary data used in the analyses were from July–August surveys, initiated in 1970, that are conducted annually by the Department of Fisheries and Oceans (DFO) of the eastern Scotian Shelf (Supplementary Fig. 1), Canada<sup>31–34</sup>. Results from a secondary, fishery-independent bottom trawl survey beginning in 1986 and conducted during the month of March were used to assess the benthic fish biomass dynamics further. We consider this a secondary survey because of its reduced geographic coverage, different statistical design, and three missing/incomplete years in comparison to the July–August survey.

The dominant, large-bodied predator species, designated the predator group, are Atlantic cod *Gadus morhua* ( $L_{\max} = 161$  cm), haddock *Melanogrammus aeglefinus* (87 cm), pollock *Pollachius virens* (112 cm), longfin hake *Urophycis chesteri* (71 cm), silver hake *Merluccius bilinearis* (73 cm), white hake *Urophycis tenuis* (142 cm), red hake *Urophycis chuss* (69 cm), redfish *Sebastes* spp. (60 cm), thorny skate *Amblyraja radiata* (120 cm), spiny dogfish *Squalus acanthias* (196 cm), Greenland halibut *Reinhardtius hippoglossoides* (82 cm), American plaice *Hippoglossoides platessoides* (76 cm), winter flounder *Pseudopleuronectes americanus* (64 cm), witch flounder *Glyptocephalus cynoglossus* (67 cm), and yellowtail flounder *Limanda ferruginea* (59 cm). These fifteen species have been commercially exploited, often in mixed fisheries, throughout the continental shelf in depths of less than 200 m. Ages are determined for Atlantic cod, haddock, pollock and silver hake; a growth model for redfish<sup>35</sup> applicable to the eastern Scotian Shelf stock and an age–length key for American plaice<sup>36</sup> permitted approximations of abundance at age for these two species. The availability of age data permitted the estimation of total mortalities (Supplementary Table 1) and growth rates (Supplementary Table 2).

The forage species group consists of three species: herring *Clupea harengus*, capelin *Mallotus villosus*, and northern sand lance *Ammodytes dubius*, which are under-sampled by the bottom trawl survey. Correction factors have been applied (capelin and sand lance by a factor of 200 and herring by 40 (ref. 37)). Commercial exploitation of these species has been relatively low (herring) or non-existent in this region. Annual anomalies (s.d. units) of individual average body weights (total biomass/total abundance) for each of the three forage species were estimated and used as indices of the temporal variation in physiological condition.

Lower trophic level data (phytoplankton colour index and zooplankton abundance time series) were obtained from Continuous Plankton Recorder (CPR) observations collected on the Scotian Shelf beginning in 1961 at a nominal sampling interval of 1 month. The CPR instrument is towed behind a vessel at about 7 m depth; plankton are collected on a 270  $\mu$ m silk mesh over a 10 nautical mile tow and stored within the recorder for later identification. Details of the sampling and analysis methods can be found on the Sir Alister Hardy Foundation website (<http://www.sahfos.ac.uk/>).

The number of CPR samples per month by year is shown in Supplementary Table 3. The series consists of two periods of data collection, 1961–1976, when there were many months without samples, and 1991–2008, when most months were sampled. There was a large data gap from 1977–1990. To compensate for missing monthly data, a time series for the period 1961–1976 was created by averaging monthly values over successive 5-year periods. For the better-sampled 1991–2008 period, 3-year averaging blocks were used. The averaging acts like a rough low-pass filter. The positions of all samples are shown in Supplementary Fig. 9.

Zooplankton data obtained from the CPR Program were grouped into large-bodied ( $\geq 2$  mm carapace length) and small-bodied ( $< 2$  mm) species. The large-bodied category comprised all *Calanus* species and their copepodite stages including *C. finmarchicus*, *C. hyperboreus*, *C. helgolandicus* and *C. glacialis*; all *Metridia*, *Euchaeta* and *Pleuromamma* species including *M. lucens*, *M. longa*, *E. marina*, *E. norvegica*, *P. abdominalis* and *P. robusta*; also included were *Candacia* sp., *C. armata*, *Heterorhabdus* sp. and *H. papilliger*, and *Neocalanus gracilis*, *Rhincalanus nasutus*, *Euchirella rostrata* and *Anomalocera patersoni*.

The small-bodied category included: all *Centropages* and *Temora* species including *C. typicus*, *C. hamatus*, *C. bradyi*, *C. chierchiai*, *T. longicornis* and *T. stylifera*; also *Pseudocalanus*, *Candacia*, *Paracalanus*, *Oncaea*, *Sapphirina*, *Lucicutia*, *Scolecithricella*, *Clausocalanus* and *Calocalanus* sp.; and *Acartia* sp., *Acartia danae*, *Corycaeus* sp., *Ctenocalanus vanus*, *Mecynocera clausi*, *Nannocalanus minor*, *Pleuromamma* sp., *Pleuromamma borealis*, *Lucicutia* and *Tortanus discaudatus*. Further details of the species identification protocols used are provided elsewhere<sup>38</sup>.

**Bottom trawl survey precision.** The precision of the biomass estimates at both the functional group and species level was evaluated by using the random, stratified survey design (27 strata) to calculate the relative standard error (rse, standard error divided by the mean and expressed as a percentage)<sup>39</sup>. For the aggregate predator biomass, the rse (averaged over 41 years) equalled 17.5% (range: 8–40%). The rse for the annual forage fish biomass was averaged over a fewer number of surveys (1993 to 2010) because, before 1993, there were very low levels of forage

fish with the majority of survey sets yielding null values ( $> 70\%$ ). The resulting rse was 32% (range: 16–59%). The rse of the annual biomasses of the individual predator species are shown in Supplementary Fig. 6. The average value for cod-like species (35%) compares favourably to surveys in other geographic regions<sup>40</sup>. The forage species biomass estimates averaged 53, 45 and 38%, for capelin, northern sand lance and herring, respectively.

We also considered the probability that the aggregate benthic fish biomass from the relatively constant period of 2000–2004 (168,000 t) was different from the 2010 value (336,000 t) represented by a linear fit to the 2004–2010 observations when the biomass had an average increase of 28,000 t per year. From the analysis above, we took 19% as a representative rse and computed the probability that the 2000–2004 value differed from the 2010 best-fit value. The probability that the 2000–2004 average value was greater than or equal to the best-fit 2010 value was 0.009.

**Commercial landings and fishery regulations.** Annual eastern Scotian Shelf landings for the benthic species were extracted from databases maintained by the Northwest Atlantic Fisheries Organization (NAFO) and the DFO. The management unit, NAFO Div. 4VW (Supplementary Fig. 1), goes beyond the continental shelf into offshore slope waters where fishing effort was concentrated for certain species, specifically silver hake. Therefore, estimation of meaningful exploitation levels based on the ratio of landings to biomass (derived from the fishery-independent trawl surveys that are constrained to the continental shelf) required the exclusion of the landings data for silver hake. No survey catchability corrections were made for any of the benthic fish species and this may result in annual ratios greater than 1. Hence, this measure of exploitation, often referred to as relative *F* (fishing mortality), is meant to serve as an index of commercial exploitation (Fig. 1c). There are no contemporary (post-2000) estimates of instantaneous fishing mortality rates because of the low levels of fishing associated with the cod and haddock fishing moratorium; for eleven species that were not aged there have never been any estimates of instantaneous fishing mortality rates.

In September of 1993 the eastern Scotian Shelf fishery for cod and haddock was closed while fisheries for flatfish, skates, redfish, silver hake (omitted) and some other minor species have remained open and, since then, total landings have averaged ~6,000 t. For silver hake, landings averaged 16,000 t during this same period and the percentage of cod by-catch associated with this directed fishery was very low (0.01%). The second leading fishery in this area was for redfish (4,500 t) and it incurred by-catches of cod of 0.8%. This information is available at <http://www.dfo-mpo.gc.ca/fm-gp/initiatives/cod-morue/strategic-mar-eng.htm>.

**Principal components analysis and regime shift detection.** The status of the eastern Scotian Shelf ecosystem was characterized by principal components analysis based on standardized anomalies (s.d. units) from five biological time series (CPR phytoplankton colour index, zooplankton (body size  $< \text{and} \geq 2$  mm), forage fish and large benthic predators) using a method previously described<sup>41</sup>. PCA1 and 2 accounted for 56.3% and 18% of the variance, respectively. We compared the time series variability of PCA1 constructed from the correlation matrices based on the data series from 1970–2010, which include data gaps for the three lower trophic levels from 1975–1991 and 2008–2009 and from the period (1970–74; 1992–2007) when all series were complete. The  $r^2$  between the two PCA1 series was 0.99. The loadings based on the 1970–2009 (1970–1974; 1992–2008) series of the five variables on PCA1 were: predators = 0.49 (0.50), forage fish complex =  $-0.44$  ( $-0.46$ ), large-bodied zooplankton = 0.43 (0.43), small-bodied zooplankton =  $-0.33$  ( $-0.32$ ), phytoplankton colour index =  $-0.52$  ( $-0.51$ ).

To objectively determine whether and when ecosystem transitions occurred, the dominant principal component axis (PCA1) from the five-variable biological analysis was subjected to a sequential *t*-test analysis of regime shift method (STARS) that identifies the magnitude and direction of shifts significant at a pre-determined  $\alpha$ -level, given both the expected cut-off length for the regime (*L*) and a parameter that designates how much of a difference from the observed mean (in s.d. units) are required before data are considered outliers<sup>29</sup> (available at <http://www.beringclimate.noaa.gov/regimes/index.html>). Choice of *L* and  $\alpha$  affect the number and duration of regime shifts<sup>29</sup>. We set *L* = 5 and  $\alpha$  = 0.01, in keeping with previous analyses using a subset of ecosystem data where shifts in Scotian Shelf cod–prey states were shown to be relatively insensitive to changes in *L* and *H* (a parameter describing the treatment of outliers) and where the relatively stringent  $\alpha$ -value was established to limit regime shift detection to only cases with strong evidence<sup>42</sup>. In our analyses, *L* = 10, *L* = 15 did not affect the timing of the biological regime shifts; setting  $\alpha$  = 0.05 only suggested one additional small regime shift in 1975 when *L* = 5.

**Carrying capacity estimation.** Estimates of the zooplankton biomass from the eastern Scotian Shelf have been compiled<sup>43</sup>; additional stations were available through the BIOCHEM database ([http://www.meds-sdmm.dfo-mpo.gc.ca/biochem/Biochem\\_e.htm](http://www.meds-sdmm.dfo-mpo.gc.ca/biochem/Biochem_e.htm)) and courtesy of C. Johnson and B. Casault. Biomass estimates were made using vertical hauls from the bottom to the surface using 200  $\mu$ m mesh nets. Observations from 1999–2008 provided reasonable spatial



coverage for the months of March, April, July and October. March and July observations were made during ecosystem fisheries surveys; April and October data, from Atlantic Zone Monitoring Program cruises, were collected on three standard, repeated sections at the northeastern (Cabot Strait Section), central (Louisbourg Section) and southwestern end (Halifax Section) of NAFO Div. 4VW (Supplementary Fig. 10).

The zooplankton biomasses are shown in Supplementary Table 4. Taking the simple average of these observations and multiplying by the surface area of the eastern shelf gives a zooplankton standing crop of  $4.8 \times 10^6$  t. Previous work<sup>44</sup> estimated a zooplankton production to biomass ratio of 6 to 9.7, and used an average value of 7 as an overall factor applied to zooplankton as a group. Another more rigorous, quantitative analysis<sup>45</sup>, estimated production to biomass (P/B) ratios for nine species, including leading components of the overall biomass, and for a single general category. Their biomass-weighted P/B ratio was 8.9. Multiplying the average biomass in Supplementary Table 4 by this factor yields an overall zooplankton production of approximately  $4.3 \times 10^7$  t. Using a 10% efficiency relating the zooplankton production to forage fish yields a rough estimate of the forage fish carrying capacity of  $4.3 \times 10^6$  t.

Similar to our calculation of carrying capacity for forage species, we estimated the annual production of phytoplankton on the eastern Scotian Shelf based on measurements from three sources<sup>46–48</sup>. The data from two sources<sup>46,47</sup> covered the months of March–August and November; the observations from the third<sup>47</sup> consisted of 13 monthly estimates from March 1991 to March 1992. All observations were reported as mg of carbon per m<sup>2</sup> per hour. We converted mg C to wet weight of phytoplankton using a factor of 42 (ref. 49), thus allowing direct comparison with our estimates of zooplankton production. Combining two sets<sup>46,47</sup> of results, we estimate the annual production of phytoplankton as  $640 \times 10^6$  t and  $240 \times 10^6$  t, respectively. These give ratios of phytoplankton to zooplankton production of 15 and 5.5 which are reasonable if the energy transfer efficiency is ~10%. One crude measure for estimating the carrying capacity for the predator complex involved taking the ratio of peak biomasses of benthic fish predators to forage fishes which is about 1:16.

**Damped harmonic oscillator calculations.** The input data set was the biomass of the three leading forage species. The period under consideration was 1994–2010. We used a temporally varying background state derived by running a 30% linear LOESS filter through the data. The filter split the variability into two components: a very low frequency component with a period estimated as 56 years from the autocorrelation function and higher frequency variability (Supplementary Fig. 8).

The anomalies indicated a simple damped harmonic oscillator-like (dSHO) variability, particularly from 1994 to 2001, shown as solid grey circles in Fig. 3. We solved for the characteristics of this variability by least squares fitting the observations to the dSHO equation:

$$A_0 + A_1 e^{(-d_1 t)} \cos(2\pi t / \tau),$$

where  $A_0$  is the mean,  $A_1$  the amplitude of the harmonic oscillation,  $d_1$  the damping rate, and  $\tau$  the periodicity of the oscillation. The solution based on optimal fitting of 1994–2001 observations, shown as a solid red line in Fig. 3, gives  $A_0 = 347,000$  t,  $A_1 = 6,100,000$  t,  $1/d_1 = 6.9$  years,  $\tau = 5.2$  years.

**Environmental indices and their relationship to the pre-collapse, collapse and post-collapse periods.** The relationship between forage fish biomass and water column temperatures was evaluated. The results showed no relationship between the two variables at 0 lag (Supplementary Fig. 5). We also lagged the forage fish biomass relative to the temperature; for lags up to 4 years, the  $r^2$  was less than 0.04. At lags of 5 years,  $r^2 = 0.21$ ; however, this is approaching the life cycle of these species and the biological import of the enhanced correlation at this lag is therefore questionable. Finally, we examined integrated temperatures up to 5 years. All  $r^2$  values were less than 0.02.

An examination of the seasonal variability of the stratification index showed that the long-term trend seen in the annual data was due mostly to the summer and secondarily to the fall series (Supplementary Fig. 11). Moreover, the magnitude of the changes during winter and spring (major bloom period) were small compared to those in summer and fall. The period, 1985–1994, of ongoing decline of the benthic fish functional group showed near-normal values of the stratification parameter during the spring bloom period. The outbreak period (1994–2001) of the forage group was weakly ( $r^2 = 0.21$ ) related (not significantly) to the spring stratification.

The seasonal variation of the phytoplankton colour index (Supplementary Fig. 12) and their correlations with the annual index, their averages and deviations

(Supplementary Table 5) indicated that the greatest contributions were from the winter and spring seasons (high correlations and variance), followed by fall (high correlation, moderate variance), and finally by summer (low correlation, lowest variance). This indicated that summer, with the greatest stratification changes, had the least impact on the annual colour index; fall, with the next greatest impact on stratification, made the second least contribution to the index. Further comparisons of the monthly colour indexes with the annual revealed that the greatest contributions were from March and April, typical months for the spring bloom on the eastern Scotian Shelf (Supplementary Table 6).

We assessed the impact of the full range of temperatures observed in the eastern Scotian Shelf using the relationship between the intrinsic rate of population growth,  $r$ , age at maturity,  $a$ , and bottom temperatures developed previously for 20 stocks of North Atlantic cod<sup>50</sup>. This analysis revealed that the estimates of  $r$  and  $a$  so derived were, during the pre-collapse period, on average +4% higher and –3% lower, respectively, than during the collapsed period and +3% higher and –2% lower, respectively, for the recovering period. We also note that the period 1987–1991, immediately preceding the cod collapse, featured some of the coldest bottom water temperatures on record. Again applying the approach detailed above, and the  $r$  and  $a$  data provided<sup>50</sup>, we estimate a maximum 10% decrease in  $r$  and a 9% increase in  $a$  relative to the collapsed and post-collapse intervals. During this cold period, fishing mortality on cod, based on the ratio of landings to biomass, averaged 65%, which is an extremely high annual rate of biomass removal. This dwarfs any possible impact of the environment as expressed through the effect of temperature on  $r$  and  $a$ .

The population dynamics and feeding ecology of grey seals on the eastern Scotian Shelf has been assessed since the early 1960s with sampling effort concentrated on the colony inhabiting Sable Island. Total population sizes were estimated from a model fit to census data on pup production (Supplementary Fig. 13).

1. Doubleday, W. G. *Studies No. 2, Manual on Groundfish Surveys in the Northwest Atlantic* (NAFO Scientific Council Studies, 1981).
2. Doubleday, W. G. & Rivard, D. Bottom trawl surveys. *Can. Spec. Pub. Fish. Aquat. Sci.* **58**, 1–273 (1981).
3. Chadwick, E. M. P. *et al.* History of annual multi-species trawl surveys on the Atlantic coast of Canada. *AZMP Bull.* **6**, 25–42 (2007).
4. Shackell, N. L. & Frank, K. T. Marine fish diversity on the Scotian Shelf, Canada. *Aquat. Conserv. Mar. Freshwat. Ecosyst.* **13**, 305–321 (2003).
5. Mayo, R. K., Gifford, V. M. & Jearld, A. Age validation of redfish, *Sebastes marinus* (L.), from the Gulf of Maine-Georges Bank Region. *J. Northwest Atl. Fish. Sci.* **2**, 13–19 (1981).
6. Fowler, G. W. & Stobo, W. T. Status of 4VW American plaice and yellowtail flounder. *DFO Can. Sci. Advis. Sec. Res. Doc.* **2000**, 1–86 (2000).
7. Bundy, A. Mass balance models of the eastern Scotian Shelf before and after the cod collapse and other ecosystem changes. *Can. Tech. Rep. Fish. Aquat. Sci.* **2520**, 1–192 (2004).
8. Richardson, A. J. *et al.* Using continuous plankton recorder data. *Prog. Oceanogr.* **68**, 27–74 (2006).
9. Smith, S. J. Analysis of data from bottom trawl surveys. *NAFO Sci. Coun. Studies* **28**, 25–53 (1996).
10. Benoit, H., Swain, D. & Chouinard, G. Using the long-term bottom-trawl survey of the Southern Gulf of St. Lawrence to understand marine fish populations and community change. *AZMP Bull.* **8**, 19–27 (2008).
11. Choi, J. S., Frank, K. T., Petrie, B. D. & Leggett, W. C. Integrated assessment of a large marine ecosystem: a case study of the devolution of the Eastern Scotian Shelf, Canada. *Oceanogr. Mar. Biol. Ann. Rev.* **42**, 47–67 (2005).
12. Litzow, M. A., Urban, J. D. & Laurel, B. J. Increased spatial variance accompanies reorganization of two continental shelf ecosystems. *Ecol. Apps* **18**, 1331–1337 (2008).
13. Harrison, G. *et al.* Optical, chemical, and biological oceanographic conditions in the Maritimes Region in 2008. *DFO Can. Sci. Advis. Sec. Res. Doc.* **2009**, 1–55 (2009).
14. Mills, E. L. & Fournier, R. O. Fish production and the marine ecosystems of the Scotian Shelf, eastern Canada. *Mar. Biol.* **54**, 101–108 (1979).
15. Tremblay, M. J. & Roff, J. C. Production estimates for Scotian Shelf copepods based on mass specific P/B ratios. *Can. J. Fish. Aquat. Sci.* **40**, 749–753 (1983).
16. Cochlan, W. Seasonal study of uptake and regeneration of nitrogen on the Scotian Shelf. *Cont. Shelf Res.* **5**, 555–577 (1986).
17. Fournier, R., Marra, J., Bohrer, R. & van Det, M. Plankton dynamics and nutrient enrichment of the Scotian Shelf. *J. Fish. Res. Bd Can.* **34**, 1004–1018 (1977).
18. Mousseau, L., Fortier, L. & Legendre, L. Annual production of fish larvae and their prey in relation to size-fractionated primary production (Scotian Shelf, NW Atlantic). *ICES J. Mar. Sci.* **55**, 44–57 (1998).
19. Cauffopé, G. & Heymans, S. Energy contents and conversion factors for sea lion's prey. *UBC Fish. Centre Res. Rep.* **13**, 225–237 (2005).
20. Myers, R. A., Mertz, G. & Fowlow, P. S. Maximum population growth rates and recovery times for Atlantic cod, *Gadus morhua*. *Fish. Bull.* **95**, 762–772 (1997).

# An ancient recipe for flood-basalt genesis

Matthew G. Jackson<sup>1</sup> & Richard W. Carlson<sup>2</sup>

Large outpourings of basaltic lava have punctuated geological time, but the mechanisms responsible for the generation of such extraordinary volumes of melt are not well known<sup>1</sup>. Recent geochemical evidence suggests that an early-formed reservoir may have survived in the Earth's mantle for about 4.5 billion years (ref. 2), and melts of this reservoir contributed to the flood basalt emplaced on Baffin Island about 60 million years ago<sup>3–5</sup>. However, the volume of this ancient mantle domain and whether it has contributed to other flood basalts is not known. Here we show that basalts from the largest volcanic event in geologic history—the Ontong Java plateau<sup>1,6,7</sup>—also exhibit the isotopic and trace element signatures proposed for the early-Earth reservoir<sup>2</sup>. Together with the Ontong Java plateau, we suggest that six of the largest volcanic events that erupted in the past 250 million years derive from the oldest terrestrial mantle reservoir. The association of these large volcanic events with an ancient primitive mantle source suggests that its unique geochemical characteristics—it is both hotter (it has greater abundances of the radioactive heat-producing elements) and more fertile than depleted mantle reservoirs—may strongly affect the generation of flood basalts.

The discovery of a surviving portion of the early-formed, homogeneous silicate Earth that existed immediately after formation of the core—referred to as primitive mantle—would place constraints on the earliest chemical evolution of the Earth and help to clarify the means by which the Earth arrived at its present geochemical state (see, for example, refs 8–10). Most models for this primitive mantle are based on the assumption that it should have relative abundances of refractory lithophile elements similar to those of carbonaceous chondrites—the presumed building blocks of the Earth<sup>11–13</sup>. However, the recent discovery of small ( $18 \pm 5$  parts per million, p.p.m.) differences in the  $^{142}\text{Nd}$  to  $^{144}\text{Nd}$  ratio ( $^{146}\text{Sm}$  decays to  $^{142}\text{Nd}$  with a half-life of 106 million years, Myr) between the Earth and chondrites suggests that the Earth's primitive mantle may not have chondritic relative abundances of the refractory lithophile elements<sup>14–16</sup>.

Instead, all known modern terrestrial mantle reservoirs may have evolved from a primitive precursor with Sm/Nd ratios 4.2–7.3% higher than that of chondrites, leading to a present-day  $^{143}\text{Nd}/^{144}\text{Nd}$  of 0.51290–0.51309, which translates to a present-day  $\epsilon^{143}\text{Nd}$  of +5.3 to +9.0, relative to the chondritic  $^{143}\text{Nd}/^{144}\text{Nd}$  ratio of 0.51263 (ref. 17); the stated uncertainty arises from the range of  $^{142}\text{Nd}/^{144}\text{Nd}$  found in modern terrestrial lavas and chondrites— $18 \pm 5$  p.p.m.—that propagates into uncertainty in the Sm/Nd ratio and hence into the value of the present-day  $^{143}\text{Nd}/^{144}\text{Nd}$  of the primitive precursor. If the expectation of chondritic relative abundances of refractory lithophile elements is removed, the only remaining signatures of primitive mantle are Pb-isotopic compositions on the geochron (the line in Pb-isotopic space defined by samples with constant U/Pb ratios over the Earth's age) and enrichment in the primordial isotope of helium,  $^3\text{He}$ , relative to the largely radiogenic isotope,  $^4\text{He}$ . All three (Nd, Pb and He) of the isotopic characteristics expected for a primitive terrestrial reservoir were identified in 62-Myr-old flood basalts emplaced on Baffin Island and West Greenland (BIWG)<sup>2–5</sup>. Employing the geochemical insights gained from BIWG, we examine some of the largest large igneous provinces

(LIPs)—volcanic provinces characterized by anomalously high rates of mantle melting that represent the largest volcanic events in the Earth's history—to determine whether they are associated with a primitive (albeit non-chondritic) mantle source.

Located in the southwestern Pacific, the Ontong Java Plateau (OJP) is the largest LIP on the Earth<sup>1,6,7</sup>. The average  $\epsilon^{143}\text{Nd}(t)$  of these lavas<sup>6,7</sup> plots close to the BIWG lavas (Fig. 1) and within the range predicted for the non-chondritic primitive mantle. Excluding the most incompatible and fluid mobile elements, the OJP lavas have relatively flat primitive-mantle-normalized trace-element patterns (Fig. 2) similar to the relatively flat patterns identified in the two highest  $^3\text{He}/^4\text{He}$  lavas from BIWG (Fig. 2). The flatness of the trace-element patterns<sup>18</sup> led Tejada *et al.* (ref. 7) to suggest that the OJP mantle is “almost” primitive, but not actually primitive because of the higher-than-chondritic  $^{143}\text{Nd}/^{144}\text{Nd}$  in the OJP lavas.

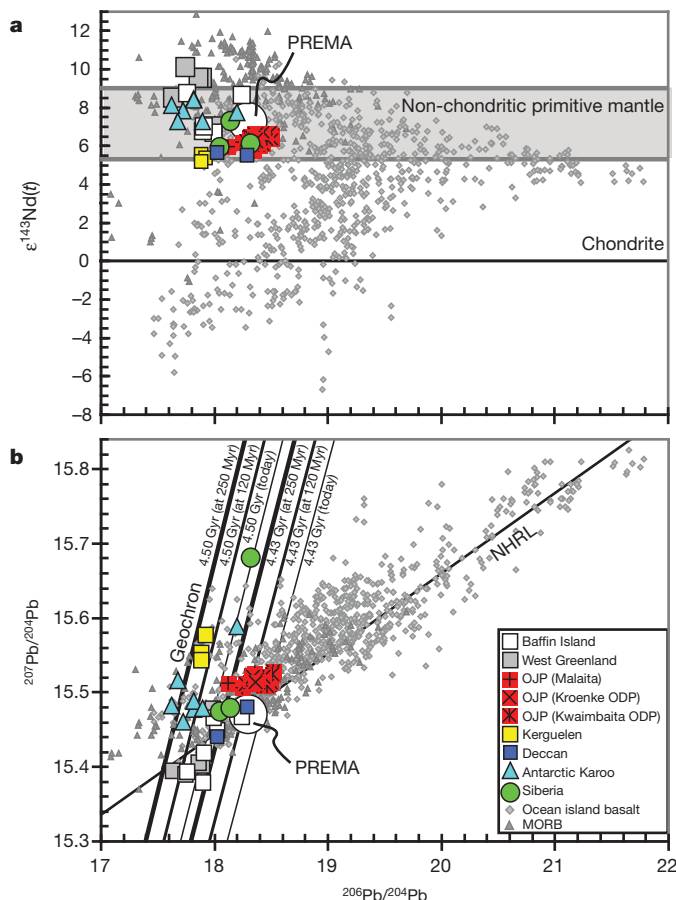
However, the discovery of a difference in  $^{142}\text{Nd}/^{144}\text{Nd}$  between modern terrestrial rocks and chondrites suggests that the  $^{143}\text{Nd}/^{144}\text{Nd}$  measured in OJP lavas overlaps with the primitive (albeit non-chondritic) terrestrial mantle. Additionally, the 120-Myr-old early-stage (Kwaimaita- and Kroenke-type lavas<sup>6,7</sup>) lavas plot near the 4.43-billion-year (Gyr)-old Pb-isotope geochron, and close to the BIWG lavas (which plot closer to the 4.5-Gyr-old geochron), an observation that is consistent with these lavas sampling an ancient mantle source (Fig. 1). Although there are minor differences in the Nd and Pb isotopic composition of the OJP and BIWG lava sources, the overlapping trace-element patterns of the two LIPs suggest an origin from compositionally similar sources whose isotopic compositions are within the range expected for an early-formed reservoir.

Owing to eruption through oceanic crust, contamination of OJP lavas by the chemically and isotopically evolved material of the continental crust does not complicate the interpretation of their mantle source to the degree seen in LIPs erupted through continental lithosphere (see Methods). In contrast, several of the largest LIPs, including the BIWG, were erupted in continental settings where assimilation of continental lithosphere can obscure the primary mantle signature of the lavas. Although continental assimilation can drive the Pb-isotopic composition of lavas towards either more or less radiogenic values, this mechanism will almost certainly lower magmatic  $^{143}\text{Nd}/^{144}\text{Nd}$  (ref. 19). To test the hypothesis that the least-contaminated LIPs that erupted in continental settings contain lavas that, like the OJP and BIWG, have Pb isotopic compositions on the geochron, we examined only the subset of LIP magmas with high  $\epsilon^{143}\text{Nd}(t)$  ( $> +5.2$ ) that fall closest to the range suggested for a non-chondritic primitive mantle. Such lavas are least likely to have suffered from assimilation of continental crust (see Methods).

The ~180-Myr-old Karoo lavas are typical of continental LIPs in that they exhibit evidence for continental crust and lithospheric mantle contamination<sup>20</sup>. However, the high-MgO lavas recently discovered in the Antarctic portion of the Karoo host high  $\epsilon^{143}\text{Nd}(t)$  ratios ( $\epsilon^{143}\text{Nd}_{180\text{ Myr}} = +7.3$  to  $+8.4$ ), and these lavas plot near the Pb geochron<sup>21</sup> (Fig. 1).

The ~251-Myr-old Siberian Traps generally have too much of a lithospheric overprint to enable us to discern primary mantle compositions<sup>22</sup>.

<sup>1</sup>Department of Earth Sciences, Boston University, 675 Commonwealth Avenue, Boston, Massachusetts 02215, USA. <sup>2</sup>Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington DC 20015, USA.

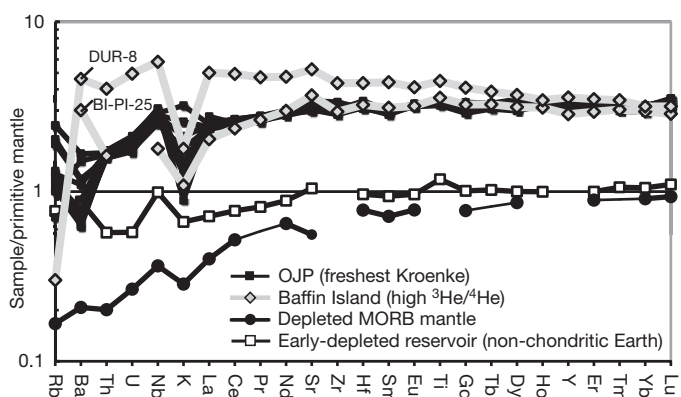


**Figure 1** | Lavas from LIPs that have Nd isotopic compositions within the range expected for a non-chondritic primitive mantle exhibit Pb isotopic compositions that plot near the geochron. In addition to having Pb-isotopic compositions and U–Pb concentrations measured on the same samples (excluding the Deccan), all samples plotted also have paired Nd-isotopic compositions and Sm–Nd concentrations. Excluding the Pb-isotopic compositions of the Deccan lavas, all isotopic compositions are age-corrected to the time of eruption and the positions of representative age-corrected geochrons are also shown. **a**, To minimize the effects of continental assimilation, only flood basalt lavas with  $\epsilon^{143}\text{Nd}(t) \geq 5.2$  are plotted. **b**, The OJP plots closest to the 4.43-Gyr geochron at the time of eruption (120 Myr). Data for the OJP include lavas from Malaita island and ODP Leg 192 drilling sites<sup>6,7</sup>; late-stage Singalo lavas are not plotted. Data sources for the other LIPs are in the text. Ocean-island-basalt and MORB data are from Georoc (<http://georoc.mpch-mainz.gwdg.de/georoc/>). The Northern Hemisphere reference line (NHRL) is shown<sup>30</sup>.  $\epsilon^{143}\text{Nd}(t) = [({}^{143}\text{Nd}/{}^{144}\text{Nd})_{\text{sample}} / ({}^{143}\text{Nd}/{}^{144}\text{Nd})_{\text{reference}} - 1] \times 10,000$ , where the isotopic ratios are calculated at the eruption age of the basalt ( $t$ ) using the chondritic reference<sup>17</sup>.

However, high-MgO meimechites<sup>22</sup> have the highest  $\epsilon^{143}\text{Nd}(t)$  ratios ( $\epsilon^{143}\text{Nd}_{251\text{ Myr}} = +5.9$  to  $+7.3$ ) associated with the Siberian Traps. These high-MgO lavas have not been affected by the continental crust contamination that is common in most Siberian Trap lavas (Fig. 1a), and have Pb isotopic compositions near the geochron.

Basalts from the Kerguelen plateau were erupted on submerged continental crust and exhibit evidence for continental crustal assimilation. Nonetheless, ~113-Myr-old Kerguelen lavas from Ocean Drilling Program (ODP) leg 120 site 749 (ref. 23) with the highest  $\epsilon^{143}\text{Nd}(t)$  values ( $\epsilon^{143}\text{Nd}_{113\text{ Myr}} = +5.2$  to  $+5.6$ ) and the lowest  $^{87}\text{Sr}/^{86}\text{Sr}$  values appear to be least affected by crustal assimilation, and their Pb isotopic compositions plot close to the 4.50-Gyr-old geochron (at 120 Myr).

Most formations of the ~65-Myr-old Deccan LIP are pervasively contaminated with crust and/or lithospheric mantle<sup>24</sup>. Lavas from the Ambenali formation have the highest  $^{143}\text{Nd}/^{144}\text{Nd}$  of the Deccan



**Figure 2** | Primitive mantle<sup>13</sup> normalized trace-element patterns of OJP lavas compared to high  $^3\text{He}/^4\text{He}$  lavas from BIWG. For the primitive mantle composition to be consistent with past usage of this term, we use the values from ref. 13, because these have served as the model, chondrite-based, primitive mantle compositions with which the scientific community is most familiar. The lavas from Baffin Island with highest  $^3\text{He}/^4\text{He}$ , BI-PI-25 and DUR8, represent the normal (N) type and enriched (E) type lavas found in the BIWG LIP, respectively. Rb was not reported for BI-PI-25, and Pb concentrations were not reported for the Baffin Island or the OJP lavas<sup>5,18</sup>. The variability in Rb and Ba (and probably K) in OJP<sup>18</sup> and BIWG<sup>2</sup> lavas is probably due to alteration, and U has been dramatically modified by alteration in BIWG sample BI-PI-25 (and is not plotted). Only Kroenke-type OJP lavas are plotted, because they have probably experienced only olivine fractionation, whereas Kwaimaita-type lavas are isotopically similar (Fig. 1) but are more evolved. All OJP and BIWG lavas are corrected for olivine fractionation to be in equilibrium with an olivine composition with a forsterite content of 92% (ref. 2). Only the freshest Kroenke-type lavas have been plotted, excluding the following samples (using the same filter as ref. 18): volcanoclastics and lavas with loss on ignition  $>0.5\%$  and/or  $\text{K}_2\text{O}/\text{P}_2\text{O}_5 > 2$ . The non-chondritic primitive mantle<sup>2</sup> (or early depleted reservoir; see Supplementary Information) and depleted MORB mantle<sup>15</sup> are included.

basalts and are thought to best represent mantle compositions. The subset of Ambenali lavas with the highest  $\epsilon^{143}\text{Nd}(t)$  ratios ( $\epsilon^{143}\text{Nd}_{65\text{ Myr}} = +5.5$  to  $+5.6$ ) have Pb isotopic ratios that straddle the geochron<sup>25</sup>.

The other test that can be applied to distinguish a primitive mantle source is high  $^3\text{He}/^4\text{He}$ . Unfortunately, owing to their age, post-eruptive radiogenic ingrowth of  $^4\text{He}$  complicates identification of mantle source  $^3\text{He}/^4\text{He}$  signatures<sup>2,3,9</sup>.

Although many of the largest LIPs host lavas that have primitive Pb and Nd isotopic compositions, a number of large LIPs erupted in the past 250 Myr do not. This could be because of the absence of a primitive mantle component in their mantle sources or could be due to pervasive continental crust assimilation by lavas with initially primitive isotopic compositions; deconvolving the effects of crustal assimilation makes it difficult to assess the abundance of the primitive component in the mantle source of continentally erupted LIPs. Critically, however, the most frequent  $^{143}\text{Nd}/^{144}\text{Nd}$  ratio in the global ocean-island basalt data set—that is, 0.5129–0.5130 (present-day  $\epsilon^{143}\text{Nd} = +5.3$  to  $+9.0$ ; see Supplementary Information)<sup>9</sup>—is identical to that predicted for a non-chondritic terrestrial mantle and is also similar to the PREMA (Prevalent Mantle; see Supplementary Information) reservoir<sup>9</sup> (Fig. 1). This narrow range of Nd-isotopic ratios also overlaps with the  $^{143}\text{Nd}/^{144}\text{Nd}$  identified in OJP and the least crustally contaminated continental LIPs (Fig. 1). If the large proportion of ocean-island-basalt lavas with present-day  $\epsilon^{143}\text{Nd}$  near  $+5.3$  to  $+9.0$  reflects a high proportion of non-chondritic primitive material in the mantle, then the early-formed reservoir must comprise a substantial portion of the modern terrestrial mantle. One problem with this interpretation is that many lavas with  $^{143}\text{Nd}/^{144}\text{Nd}$  near 0.5130 (present-day  $\epsilon^{143}\text{Nd} = +7.2$ ) do not have Pb isotopic compositions that plot on the geochron (see Supplementary Information). This may be due to



incorporation of recycled oceanic crust into the primitive reservoir, which would cause the mixture to be displaced to the right of the geochron<sup>2</sup>, owing to the high Pb concentrations and high U/Pb ratios in recycled oceanic crust<sup>26</sup>. The high Pb/Nd concentration ratios in subducted oceanic crust relative to other mantle reservoirs, however, would lead to only slightly lower  $^{143}\text{Nd}/^{144}\text{Nd}$  in the mixture.

If the Earth's primitive mantle currently has a  $^{143}\text{Nd}/^{144}\text{Nd}$  of 0.5130 (present-day  $\epsilon^{143}\text{Nd} = +7.2$ ), then mass-balance calculations suggest that only 45–60% of the mantle was depleted to generate the chemically complementary continental crust and depleted mid-ocean-ridge basalt (MORB) mantle<sup>2</sup>. Recent dynamic convection models of the Earth's mantle suggest that, over the age of the Earth, 10–25% of the Earth's convecting mantle has never melted<sup>27</sup>. Mantle stirring and mixing is likely to have contaminated much of the surviving primitive reservoir with recycled material. This is consistent with the hypothesis that the ocean-island-basalt source mantle is largely composed of non-chondritic primitive mantle material that has been slightly modified by the incorporation of recycled material. Therefore, the present-day mass of surviving primitive material that has escaped melting and incorporation of recycled material may be small. The fact that the Earth's major LIPs have Nd and Pb isotopic compositions close to those expected for the non-chondritic primitive reservoir suggests that these LIPs may preferentially sample the uncontaminated portions of this reservoir.

A significant fraction of the world's largest volcanic events tap a mantle source that is similar to the primitive, non-chondritic mantle reservoir discovered in high  $^3\text{He}/^4\text{He}$  lavas from BIWG, and the geochemical characteristics of a non-chondritic mantle may provide insights regarding its presence in LIPs. A (non-chondritic) primitive mantle has higher concentrations of the radioactive, heat-producing elements U, Th and K (0.012, 0.046 and 159 p.p.m., respectively<sup>2</sup>) than does depleted MORB mantle (0.0054, 0.016 and 68.4 p.p.m.; ref. 15), though still lower than for estimates of primitive mantle based on chondritic models (0.0203, 0.0795 and 240 p.p.m.; ref. 13). This higher heat-generating capacity would make the less-depleted mantle more prone to mantle upwelling and plume generation and also cause it to melt to a greater degree. Additionally, the bulk composition of a non-chondritic mantle would be more fertile and fusible than depleted MORB mantle. Therefore, under normal melting conditions, it would undergo greater degrees of melting. These two mechanisms—a hotter and more fusible mantle source—may work together to produce the extraordinarily large volumes of melt observed at LIPs.

The current distribution of non-chondritic primitive material in the mantle is not known, but a recently discovered palaeo-geographic relationship between Phanerozoic LIPs and the large low-shear-wave-velocity provinces (LLSVPs) in the deepest mantle might provide a clue. Torsvik *et al.* (ref. 28) showed that, using a reasonable plate reconstruction, most LIPs that erupted during the past 320 Myr plot directly over one of the two hot and dense LLSVPs, one beneath Africa and one beneath the Pacific. They suggest that the two near-antipodal LLSVPs have remained stable features in the mantle for up to 540 Myr. The geographic association of LIPs at the surface with LLSVPs at depth suggest that the mantle domain represented by the LLSVPs may be the source of the LIPs. If LIPs sample a primitive non-chondritic mantle source, then the results of Torsvik *et al.* suggest that the LLSVPs are clear candidates for hosting this ancient mantle source. At about 2% of the mantle's mass<sup>29</sup>, the two LLSVPs are sufficiently large to source all LIP volcanism over the age of the Earth (see Supplementary Information), an observation that highlights their potential as reservoirs of early primitive material in the Earth.

## METHODS SUMMARY

OJP lavas were erupted in an oceanic environment where assimilation of continental crust does not complicate the geochemical signatures of the lavas, making it possible clearly to identify their mantle source compositions. Although the OJP lavas potentially could have been contaminated by the oceanic crust through which they

erupted, the relatively small difference in isotopic composition among the OJP lavas, and the low Nd and Pb concentrations of typical basaltic oceanic crust, would cause only minor isotopic shifts in the pre-contamination magmas, given reasonable amounts of contamination. Assimilation of oceanic pelagic sediments could more dramatically alter the isotopic composition of primitive OJP lavas, but such materials tend to have low  $^{143}\text{Nd}/^{144}\text{Nd}$  and Pb isotope compositions plotting well to the right of the geochron. The OJP lavas display no trends consistent with the incorporation of pelagic sediments. In contrast, flood basalts erupted in continental settings are prone to contamination by assimilation of continental crust. In spite of being erupted in a continental setting, the subset of BIWG flood basalt lavas examined by ref. 2 exhibit little evidence for contamination by continental crust. In the Methods section, we examine a subset of lavas from each of four additional continental flood basalt provinces—the Antarctic portion of the Karoo flood basalt, the Siberian Traps, the Kerguelen Plateau and the Deccan—that, like the OJP and a subset of the BIWG, have suffered minimal crustal contamination and exhibit mantle source signatures predicted for a (non-chondritic) primitive mantle signature. Each of the four additional continental flood basalts have been extensively studied by others (see Methods), and the geochemical signatures associated with continental assimilation have also been thoroughly explored: the lavas that have suffered from the least continental crust assimilation also have the highest  $^{143}\text{Nd}/^{144}\text{Nd}$  ( $\epsilon^{143}\text{Nd}$ ) ratios, and these ratios cluster near the range predicted for a non-chondritic primitive mantle.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 14 February; accepted 22 June 2011.**

**Published online 27 July 2011.**

- Coffin, M. C. & Eldholm, O. Large igneous provinces: crustal structure, dimensions, and external consequences. *Rev. Geophys.* **32**, 1–36 (1994).
- Jackson, M. G. *et al.* Evidence for the survival of the oldest terrestrial mantle reservoir. *Nature* **466**, 853–856 (2010).
- Graham, D. W. *et al.* Helium isotope composition of the early Iceland mantle plume inferred from the Tertiary picrites of West Greenland. *Earth Planet. Sci. Lett.* **160**, 241–255 (1998).
- Stuart, F. M., Lass-Evans, S., Fitton, J. G. & Ellam, R. M. High  $^3\text{He}/^4\text{He}$  ratios in picritic basalts from Baffin Island and the role of a mixed reservoir in mantle plumes. *Nature* **424**, 57–59 (2003).
- Starkey, N. A. *et al.* Helium isotopes in early Iceland plume picrites: Constraints on the composition of high  $^3\text{He}/^4\text{He}$  mantle. *Earth Planet. Sci. Lett.* **277**, 91–100 (2009).
- Tejada, M. L. G., Mahoney, J. J., Neal, C. R., Duncan, R. A. & Petterson, M. G. Basement geochemistry and geochronology of central Malaita, Solomon Islands, with implications for the origin and evolution of the Ontong Java Plateau. *J. Petrol.* **43**, 449–484 (2002).
- Tejada, M. L. G. *et al.* in *Origin and Evolution of the Ontong Java Plateau* (eds Fitton, J. G., Mahoney, J. J., Wallace, P. J. & Saunders, A. D.) 133–150 (Geological Society of London Special Publication 229, 2004).
- Jacobsen, S. B. & Wasserburg, G. J. The mean age of mantle and crustal reservoirs. *J. Geophys. Res.* **84**, 7411–7427 (1979).
- Zindler, A. & Hart, S. R. Chemical geodynamics. *Annu. Rev. Earth Planet. Sci.* **14**, 493–571 (1986).
- DePaolo, D. J. Crustal growth and mantle evolution: inferences from models of element transport and Nd and Sr isotopes. *Geochim. Cosmochim. Acta* **44**, 1185–1196 (1980).
- Jagoutz, E. *et al.* The abundances of major, minor and trace elements in the Earth's mantle as derived from primitive ultramafic nodules. *Proc. Lunar Planet. Sci. Conf.* **X**, 2031–2050 (1979).
- Palme, H. & O'Neill, H. S. C. in *The Mantle and Core: Treatise on Geochemistry* (ed. Carlson, R. W.) Vol. 2, 1–38 (Elsevier, 2003).
- McDonough, W. F. & Sun, S. S. The composition of the Earth. *Chem. Geol.* **120**, 223–253 (1995).
- Boyett, M. & Carlson, R. W.  $^{142}\text{Nd}$  evidence for early (>4.53 Ga) global differentiation of the silicate Earth. *Science* **309**, 576–581 (2005).
- Boyett, M. & Carlson, R. W. A new geochemical model for the Earth's mantle inferred from  $^{146}\text{Sm}$ – $^{142}\text{Nd}$  systematic. *Earth Planet. Sci. Lett.* **250**, 254–268 (2006).
- Caro, G. & Bourdon, B. Non-chondritic Sm/Nd ratio in the terrestrial planets: consequences for the geochemical evolution of the mantle–crust system. *Geochim. Cosmochim. Acta* **74**, 3333–3349 (2010).
- Bouvier, A., Vervoort, J. D. & Patchett, P. J. The Lu–Hf and Sm–Nd isotopic composition of CHUR: constraints from unequilibrated chondrites and implications for the bulk composition of terrestrial planets. *Earth Planet. Sci. Lett.* **273**, 48–57 (2008).
- Fitton, J. G. & Goddard, M. in *Origin and Evolution of the Ontong Java Plateau* (eds Fitton, J. G., Mahoney, J. J., Wallace, P. J. & Saunders, A. D.) 151–178 (Geological Society of London Special Publication 229, 2004).
- Carlson, R. W. Physical and chemical evidence on the cause and source characteristics of flood basalt volcanism. *Aust. J. Earth Sci.* **38**, 525–544 (1991).
- Hawkesworth, C. J., Marsh, J. S., Duncan, A. R., Erlank, A. J. & Norry, M. J. in *Petrogenesis of the Volcanic Rocks of the Karoo Province* (ed. Erlank, A. J.) 341–354 (Geological Society of South Africa Special Publication, 1984).

21. Heinonen, J. S., Carlson, R. W. & Luttinen, A. V. Isotopic (Sr, Nd, Pb, and Os) composition of highly magnesian dikes of Vestfjella, western Dronning Maud Land, Antarctica: a key to the origins of the Jurassic Karoo large igneous province. *Chem. Geol.* **277**, 227–244 (2010).
22. Carlson, R. W., Czamanske, G., Fedorenko, V. & Ilupin, I. A comparison of Siberian meimechites and kimberlites: implications for the source of high-Mg alkalic magmas and flood basalts. *Geochem. Geophys. Geosyst.* **7**, Q11014 (2006).
23. Frey, F. A., Weis, D., Borisova, A. Y. & Xu, G. Involvement of continental crust in the formation of the Cretaceous Kerguelen Plateau: new perspectives from ODP Leg 120 Sites. *J. Petrol.* **43**, 1207–1239 (2002).
24. Peng, Z. X., Mahoney, J. J., Hooper, P. R., Harris, C. & Beane, J. E. A role for lower continental crust in flood basalt genesis? Isotopic and incompatible element study of the lower six formations of the western Deccan Traps. *Geochim. Cosmochim. Acta* **58**, 267–288 (1994).
25. Lightfoot, P. C., Hawkesworth, C. J., Devey, C. W., Rogers, N. W. & van Calsteren, P. W. C. Source and differentiation of Deccan Trap Lavas: implications of geochemical and mineral chemical variations. *J. Petrol.* **31**, 1165–1200 (1990).
26. Hofmann, A. W. & White, W. M. Mantle plumes from ancient oceanic crust. *Earth Planet. Sci. Lett.* **57**, 421–436 (1982).
27. Brandenburg, J. P., Hauri, E. H., van Keken, P. E. & Ballentine, C. J. A multiple-system study of the geochemical evolution of the mantle with force-balanced plates and thermochemical effects. *Earth Planet. Sci. Lett.* **276**, 1–13 (2008).
28. Torsvik, T. H., Burke, K., Steinberger, B., Webb, S. J. & Ashwal, L. D. Diamonds sampled by plumes from the core–mantle boundary. *Nature* **466**, 352–355 (2010).
29. Burke, K. Plate tectonics, the Wilson Cycle, and mantle plumes: geodynamics from the top. *Annu. Rev. Earth Planet. Sci.* **39**, 1–29 (2011).
30. Hart, S. R. A large-scale isotope anomaly in the southern hemisphere mantle. *Nature* **309**, 753–757 (1984).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank J. Day, S. Hart, N. Shimizu, S. Shirey, J. Mahoney and M. Kurz for discussions and B. Hanan and G. Fitton for their detailed review comments. M.J. acknowledges Boston University start-up funds and the Ocean Sciences Section of the National Science Foundation that supported this work.

**Author Contributions** M.J. and R.C. contributed equally to the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.G.J. ([jacksonm@bu.edu](mailto:jacksonm@bu.edu)).

## METHODS

To establish a baseline for the composition of melts sampling a (non-chondritic) primitive mantle reservoir, this Letter relies on the conclusions drawn from the OJP data set, in large part to avoid the issue of the sensitivity of continental flood basalts to contamination by continental crust during emplacement and eruption. The OJP was erupted in an oceanic environment, and there is no evidence for continental fragments in the OJP. Although we cannot rule out small degrees of contamination from the oceanic lithosphere, the Pb-isotopic composition of OJP lavas is not at all consistent with a strong oceanic crust or sediment overprint. The tight clustering of the OJP Pb-isotopic data—which were measured on samples collected from Malaita and multiple drill core locations scattered over hundreds of kilometres across the plateau—is remarkable and argues against any systematic offset of the data caused by oceanic crust or sediment contamination. An important point regarding the OJP lavas is that we did not filter the data on the basis of any isotopic parameter. Figure 1 shows all available OJP lavas (excluding late-stage Singalo lavas) that have both paired U–Pb concentrations and Pb-isotopic measurements and paired Sm–Nd concentrations and Nd-isotopic measurements.

In contrast, a specific isotopic filter was applied to the continental flood basalts, because the isotopic compositions of such lavas are often overprinted by continental crust. In selecting samples with the most radiogenic Nd-isotopic compositions, we are selecting samples with the lowest crustal input and it is those lavas that have Pb-isotopic compositions that plot close to the geochron. Although we cannot rule out the possibility that these basalts obtained their Pb-isotopic compositions through a haphazard mixture of continental crust and variable mantle sources, it is worth noting that all of the samples shown in Fig. 1—that were selected based solely on Nd-isotopic compositional characteristics—also plot exactly where expected in Pb-isotopic space for a melt of ancient primitive mantle. Below, we present additional geochemical evidence from the literature that supports our contention that continental flood basalts—including BIWG, the Kerguelen Plateau, Siberian Traps, Deccan Traps and the Antarctic portion of the Karoo—with the highest  $\epsilon^{143}\text{Nd}/^{144}\text{Nd}$  have suffered from the least crustal contamination and best represent mantle source compositions.

**BIWG flood basalt lavas.** Continental crust contamination had already been ruled out for a suite of lavas from BIWG examined by ref. 2. Larsen and Pederson (ref. 31) used trace-element indicators sensitive to continental crust assimilation to show that West Greenland lavas with the least crustal contamination have radiogenic  $\epsilon^{143}\text{Nd}(t)$  values and plot in a specific region of Pb-isotopic space. An evaluation of this data by ref. 2 showed that the region of Pb-isotopic space occupied by the least contaminated Baffin Island lavas corresponds with the 4.55–4.45-Gyr terrestrial geochrons.

**Kerguelen Plateau basalts.** Although originally viewed as an oceanic plateau, continental fragments have been identified in the Kerguelen Plateau using seismic methods (ref. 32). Clasts of garnet-biotite gneiss recovered from Elan Bank contain zircons with ages ranging from 534 to 2,547 Myr (ref. 33), clearly indicative of an old crustal basement beneath portions of the Kerguelen Plateau. Additionally, studies in the area have found geochemical evidence for continental assimilation in Kerguelen basalts, and workers in this area have spent considerable efforts showing which basalts are and are not contaminated in this province<sup>23,34–37</sup>. For example, Kerguelen lavas that have been contaminated with continental crust tend to have low Nb/La (for example, ODP leg 120 site 747, ODP leg 183 site 1137) or high  $^{87}\text{Sr}/^{86}\text{Sr}$  ( $>0.7045$ ; for example, ODP leg 119 site 738, ODP leg 120 sites 747 and 750, and ODP leg 183 site 1137). In contrast, a subset of the lavas from ODP leg 120 site 749 have low  $^{87}\text{Sr}/^{86}\text{Sr}$  ( $<0.704$ ) and Nb/La ratios (0.9–1.0) that are similar to oceanic lavas and low  $^{87}\text{Sr}/^{86}\text{Sr}$  ratios<sup>23</sup>. Site 749 lavas are divided into two distinct groups: group 1 has lower Nb/La (0.80–0.85) and higher  $^{87}\text{Sr}/^{86}\text{Sr}$  ( $>0.704$ ) than group 2, which has higher Nb/La (0.9–1.0) and lower  $^{87}\text{Sr}/^{86}\text{Sr}$ . The site 749 lavas from group 1, with the lower Nb/La and higher  $^{87}\text{Sr}/^{86}\text{Sr}$ , are interpreted to have suffered from more assimilation of continental crust than the group 2 lavas. We note that the group 2 lavas from site 749 basalts, which we consider to be the least contaminated by continental crust, have the highest  $\epsilon^{143}\text{Nd}(t)$  in the Kerguelen Plateau suite, an observation that supports our contention that continental LIPs with the most radiogenic Nd-isotopic compositions best represent the mantle source in geochemically complex continental flood basalt provinces. These lavas also have Pb-isotopic compositions that plot close to the geochron.

**Deccan flood basalt lavas.** In the Deccan, the Ambenali formation flows are thought to best represent mantle compositions<sup>24,38,39</sup>. Several studies have documented the nature of crustal contamination in the Deccan and show that the least-contaminated members of the Ambenali are closest to mantle source compositions<sup>24,38,39</sup>. For example, ref. 38 used Ba/Y and Ba/Zr (where both ratios are high in continental crust) to identify lavas that have suffered from crustal contamination, and they found that lavas with the lowest Ba/Y and Ba/Zr are found in the Ambenali formation. Additionally, ref. 39 used Ba/Ti to evaluate continental assimilation (where high Ba/Ti is high in continental crust), and they found Ambenali lavas to have low Ba/Ti ratios. Ref. 39 also considered radiogenic Sr isotopic ratios to be a signature of crustally contaminated Deccan lavas, and although Ambenali lavas tend to have the lowest  $^{87}\text{Sr}/^{86}\text{Sr}$  in the Deccan, some of the Ambenali lavas have elevated  $^{87}\text{Sr}/^{86}\text{Sr}$ —with correspondingly low  $\epsilon^{143}\text{Nd}(t)$ —that are considered to be crustally contaminated. Like ref. 39, we exclude such lavas as having a mantle source composition. In summary, the least-contaminated units of the Ambenali formation, which is considered to be the least contaminated in the Deccan, also have the highest  $\epsilon^{143}\text{Nd}(t)$ , and these lavas have Pb-isotopic compositions that plot close to the geochron.

**Siberian flood basalt.** The ~251-Myr-old Siberian Traps are generally too contaminated by continental crust assimilation to allow us to evaluate primary mantle compositions<sup>22,40</sup>. However, a recent study showed that the Siberian meimechites have the highest  $\epsilon^{143}\text{Nd}(t)$  of any Siberian flood basalt and have trace-element characteristics inconsistent with crustal contamination<sup>22</sup>. Together with having the highest  $\epsilon^{143}\text{Nd}(t)$ , the meimechites also have the lowest Th/Ta in the flood basalt province. This is in contrast to the crustally contaminated Siberian flood basalt lavas that have high Th/Ta (continental crust has high Th/Ta) coupled with low  $\epsilon^{143}\text{Nd}(t)$  (ref. 22). The Siberian meimechites also have Os-isotopic compositions inconsistent with any significant crustal input<sup>22</sup>. Like the Deccan and Kerguelen flood basalts, the least crustally contaminated Siberian flood basalt lavas have the highest  $\epsilon^{143}\text{Nd}(t)$ , and these lavas have Pb-isotopic compositions that plot close to the geochron.

**Antarctic Karoo flood basalts.** Using a data set similar to the Siberian lavas from ref. 22, similar arguments were used to argue for a lack of contamination in the picritic lavas in the Antarctic portion of the Karoo flood basalt province<sup>21</sup>. The lavas that are considered to be least contaminated by continental crust exhibit Os-isotopic signatures that are thought to reflect mantle signatures (age-corrected  $^{188}\text{Os}/^{188}\text{Os} = 0.1257\text{--}0.1286$ ), while the crustally contaminated lavas have higher  $^{187}\text{Os}/^{188}\text{Os}$  (0.1297–0.1426), consistent with continental crust assimilation. The crustally contaminated lavas also have higher  $^{87}\text{Sr}/^{86}\text{Sr}$  ( $\geq 0.7036$  after age correction) than the least crustally contaminated lavas ( $<0.7036$ ). The least-contaminated Antarctic Karoo lavas also have the highest  $\epsilon^{143}\text{Nd}(t)$ , and their Pb-isotopic compositions plot close to the geochron.

- Larsen, L. M. & Pedersen, A. K. Petrology of the Paleocene picrites and flood basalts on Disko and Nuussuaq, West Greenland. *J. Petrol.* **50**, 1667–1711 (2009).
- Operto, S. & Charvis, P. Deep structure of the southern Kerguelen Plateau (southern Indian Ocean) from ocean bottom seismometer wide-angle seismic data. *J. Geophys. Res.* **101**, 25077–25103 (1996).
- Nicolaysen, K. *et al.* Provenance of Proterozoic garnet-biotite gneiss recovered from Elan Bank, Kerguelen Plateau, southern Indian Ocean. *Geology* **29**, 235–238 (2001).
- Ingle, S., Weis, D., Doucet, S. & Mattioli, N. Hf isotope constraints on mantle sources and shallow-level contaminants during Kerguelen hotspot activity since ~120 Ma. *Geochim. Geophys. Geosyst.* **4**, (2003).
- Mahoney, J. *et al.* Geochemical characteristics of lavas from Broken Ridge, the Naturaliste Plateau and Southernmost Kerguelen Plateau: early volcanism of the Kerguelen hotspots. *Chem. Geol.* **120**, 315–345 (1995).
- Ingle, S., Weis, D., Scoates, J. S. & Frey, F. A. Relationship between the early Kerguelen plume and continental flood basalts of the paleo-Eastern Gondwanan margins. *Earth Planet. Sci. Lett.* **197**, 35–50 (2002).
- Storey, M. *et al.* Lower Cretaceous volcanic rocks on continental margins and their relationship to the Kerguelen Plateau. *Proc. ODP Sci. Res.* **120**, 33–53 (1992).
- Lightfoot, P. & Hawkesworth, C. Origin of Deccan Trap lavas: evidence from combined trace element and Sr-, Nd- and Pb-isotope studies. *Earth Planet. Sci. Lett.* **91**, 89–104 (1988).
- Peng, Z. X. & Mahoney, J. J. Drillhole lavas from the northwestern Deccan Traps, and the evolution of R union hotspot mantle. *Earth Planet. Sci. Lett.* **134**, 169–185 (1995).
- Fedorenko, V. A. *et al.* Petrogenesis of the flood-basalt sequence at Noril'sk, North Central Siberia. *Int. Geol. Rev.* **38**, 99–135 (1996).



# Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma

Ryan D. Morin<sup>1\*</sup>, Maria Mendez-Lago<sup>1\*</sup>, Andrew J. Mungall<sup>1</sup>, Rodrigo Goya<sup>1</sup>, Karen L. Mungall<sup>1</sup>, Richard D. Corbett<sup>1</sup>, Nathalie A. Johnson<sup>2</sup>, Tessa M. Severson<sup>1</sup>, Readman Chiu<sup>1</sup>, Matthew Field<sup>1</sup>, Shaun Jackman<sup>1</sup>, Martin Krzywinski<sup>1</sup>, David W. Scott<sup>2</sup>, Diane L. Trinh<sup>1</sup>, Jessica Tamura-Wells<sup>1</sup>, Sa Li<sup>1</sup>, Marlo R. Firme<sup>1</sup>, Sanja Rogic<sup>2</sup>, Malachi Griffith<sup>1</sup>, Susanna Chan<sup>1</sup>, Oleksandr Yakovenko<sup>1</sup>, Irmtraud M. Meyer<sup>3</sup>, Eric Y. Zhao<sup>1</sup>, Duane Smailus<sup>1</sup>, Michelle Moksa<sup>1</sup>, Suganthi Chittaranjan<sup>1</sup>, Lisa Rimsza<sup>4</sup>, Angela Brooks-Wilson<sup>1,5</sup>, John J. Spinelli<sup>6,7</sup>, Susana Ben-Neriah<sup>2</sup>, Barbara Meissner<sup>2</sup>, Bruce Woolcock<sup>2</sup>, Merrill Boyle<sup>2</sup>, Helen McDonald<sup>1</sup>, Angela Tam<sup>1</sup>, Yongjun Zhao<sup>1</sup>, Allen Delaney<sup>1</sup>, Thomas Zeng<sup>1</sup>, Kane Tse<sup>1</sup>, Yaron Butterfield<sup>1</sup>, Inanç Birol<sup>1</sup>, Rob Holt<sup>1</sup>, Jacqueline Schein<sup>1</sup>, Douglas E. Horsman<sup>2</sup>, Richard Moore<sup>1</sup>, Steven J. M. Jones<sup>1</sup>, Joseph M. Connors<sup>2</sup>, Martin Hirst<sup>1</sup>, Randy D. Gascoyne<sup>2,8</sup> & Marco A. Marra<sup>1,9</sup>

**Follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL) are the two most common non-Hodgkin lymphomas (NHLs). Here we sequenced tumour and matched normal DNA from 13 DLBCL cases and one FL case to identify genes with mutations in B-cell NHL. We analysed RNA-seq data from these and another 113 NHLs to identify genes with candidate mutations, and then re-sequenced tumour and matched normal DNA from these cases to confirm 109 genes with multiple somatic mutations. Genes with roles in histone modification were frequent targets of somatic mutation. For example, 32% of DLBCL and 89% of FL cases had somatic mutations in *MLL2*, which encodes a histone methyltransferase, and 11.4% and 13.4% of DLBCL and FL cases, respectively, had mutations in *MEF2B*, a calcium-regulated gene that cooperates with CREBBP and EP300 in acetylating histones. Our analysis suggests a previously unappreciated disruption of chromatin biology in lymphomagenesis.**

Non-Hodgkin lymphomas (NHLs) are cancers of B, T or natural killer lymphocytes. The two most common types of NHL, follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL), together comprise 60% of new B-cell NHL diagnoses each year in North America<sup>1</sup>. FL is an indolent and typically incurable disease characterized by clinical and genetic heterogeneity. DLBCL is aggressive and likewise heterogeneous, comprising at least two distinct subtypes that respond differently to standard treatments. Both FL and the germinal centre B-cell (GCB) cell of origin (COO) subtype of DLBCL derive from germinal centre B cells, whereas the activated B-cell (ABC) variety, which has a more aggressive clinical course, is thought to originate from B cells that have exited, or are poised to exit, the germinal centre<sup>2</sup>. Current knowledge of the specific genetic events leading to DLBCL and FL is limited to the presence of a few recurrent genetic abnormalities<sup>2</sup>. For example, 85–90% of FL and 30–40% of GCB DLBCL cases<sup>3,4</sup> harbour t(14;18)(q32;q21), which results in deregulated expression of the *BCL2* oncoprotein. Other genetic abnormalities unique to GCB DLBCL include amplification of the *c-REL* gene and of the miR-17-92 microRNA cluster<sup>5</sup>. In contrast to GCB cases, 24% of ABC DLBCLs harbour structural alterations or inactivating mutations affecting *PRDM1*, which is involved in differentiation of GCB cells into antibody-secreting plasma cells<sup>6</sup>. ABC-specific mutations also affect genes regulating NF- $\kappa$ B signalling<sup>7,8,9</sup>, with *TNFAIP3* (also known as *A20*) and *MYD88* (ref. 10) the most abundantly mutated in 24% and 39% of cases, respectively. To enhance our understanding of the genetic architecture of B-cell NHL, we undertook a study to (1) identify somatic mutations and

(2) determine the prevalence, expression and focal recurrence of mutations in FL and DLBCL. Using strategies and techniques applied to cancer genome and transcriptome characterization by ourselves and others<sup>11,12,13</sup>, we sequenced tumour DNA and/or RNA from 117 tumour samples and 10 cell lines (Supplementary Tables 1 and 2) and identified 651 genes (Supplementary Figure 1) with evidence of somatic mutation in B-cell NHL. After validation, we showed that 109 genes were somatically mutated in two or more NHL cases. We further characterized the frequency and nature of mutations within *MLL2* and *MEF2B*, which were among the most frequently mutated genes with no previously known role in lymphoma.

## Identification of recurrently mutated genes

We sequenced the genomes or exomes of 14 NHL cases, all with matched constitutional DNA sequenced to comparable depths (Supplementary Tables 1 and 2). After screening for single nucleotide variants followed by subtraction of known polymorphisms and visual inspection of the sequence read alignments, we identified 717 non-synonymous variants (coding single nucleotide variants; cSNVs) affecting 651 genes (Supplementary Figure 1 and Supplementary Methods). We identified between 20 and 135 cSNVs in each of these genomes. Only 25 of the 651 genes with cSNVs were represented in the cancer gene census (December 2010 release)<sup>14</sup>.

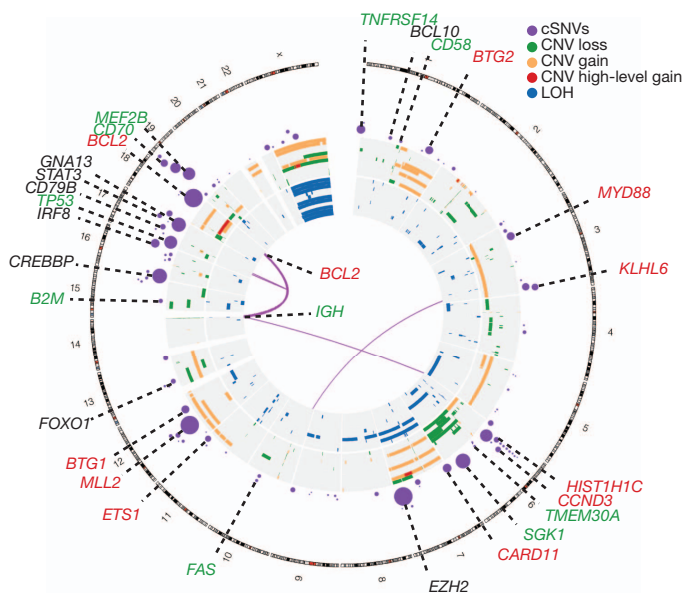
We performed RNA sequencing (RNA-seq) on these 14 NHL cases and an expanded set of 113 samples comprising 83 DLBCL, 12 FL and 8 B-cell NHL cases with other histologies and 10 DLBCL-derived cell lines (Supplementary Table 2). We analysed these data to identify

<sup>1</sup>Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z 1L3, Canada. <sup>2</sup>Centre for Lymphoid Cancer, BC Cancer Agency, Vancouver, British Columbia V5Z 1L3, Canada. <sup>3</sup>Centre for High-throughput Biology, Department of Computer Science, Vancouver, British Columbia V6T 1Z4, Canada. <sup>4</sup>Department of Pathology, University of Arizona, Tucson, Arizona 85724, USA. <sup>5</sup>Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada. <sup>6</sup>Cancer Control Research, BC Cancer Agency, Vancouver, British Columbia V5Z 1L3, Canada. <sup>7</sup>School of Population and Public Health, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada. <sup>8</sup>Department of Pathology, University of British Columbia, Vancouver, British Columbia V6T 2B5, Canada. <sup>9</sup>Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia V6H 3N1, Canada.

\*These authors contributed equally to this work.

novel fusion transcripts (Supplementary Table 3) and cSNVs (Fig. 1). We identified 240 genes with at least one cSNV in a genome/exome or an RNA-seq 'mutation hot spot' (see later), and with cSNVs in at least three cases in total (Supplementary Table 4). We selected cSNVs from each of these 240 genes for re-sequencing to confirm their somatic status. We did not re-sequence genes with previously documented mutations in lymphoma (for example, *CD79B*, *BCL2*). We confirmed the somatic status of 543 cSNVs in 317 genes, with 109 genes having at least two confirmed somatic mutations (Supplementary Tables 4 and 5). Of the successfully re-sequenced cSNVs predicted from the genomes, 171 (94.5%) were confirmed somatic, 7 were false calls and 3 were present in the germ line. These 109 recurrently mutated genes were significantly enriched for genes implicated in lymphocyte activation ( $P = 8.3 \times 10^{-4}$ ; for example, *STAT6*, *BCL10*), lymphocyte differentiation ( $P = 3.5 \times 10^{-3}$ ; for example, *CARD11*), and regulation of apoptosis ( $P = 1.9 \times 10^{-3}$ ; for example, *BTG1*, *BTG2*). Also significantly enriched were genes linked to transcriptional regulation ( $P = 5.4 \times 10^{-4}$ ; for example, *TP53*) and genes involved in methylation ( $P = 2.2 \times 10^{-4}$ ) and acetylation ( $P = 1.2 \times 10^{-2}$ ), including histone methyltransferase (HMT) and acetyltransferase (HAT) enzymes known previously to be mutated in lymphoma (for example, *EZH2* (ref. 13) and *CREBBP* (ref. 15); Supplementary Methods).

Mutation hot spots can result from mutations at sites under strong selective pressure and we have previously identified such sites using RNA-seq data<sup>13</sup>. We searched our RNA-seq data for genes with mutation hot spots, and identified 10 genes that were not mutated in the 14 genomes (*PIM1*, *FOXO1*, *CCND3*, *TP53*, *IRF4*, *BTG2*, *CD79B*, *BCL7A*, *IKZF3* and *B2M*), of which five (*FOXO1*, *CCND3*, *BTG2*,



**Figure 1 | Genome-wide visualization of somatic mutation targets in NHL.** Overview of structural rearrangements and copy number variations (CNVs) in the 11 DLBCL genomes and cSNVs in the 109 recurrently mutated genes identified in our analysis. Inner arcs represent somatic fusion transcripts identified in at least one of the 11 genomes. The CNVs and LOH detected in each of the 11 DLBCL tumour/normal pairs are displayed on the concentric sets of rings. The inner 11 rings show regions of enhanced homozygosity plotted with blue (interpreted as LOH). The outer 11 rings show somatic CNVs. Purple circles indicate the position of genes with at least two confirmed somatic mutations with circle diameter proportional to the number of cases with cSNVs detected in that gene. Circles representing the genes with significant evidence for positive selection are labelled. Coincidence between recurrently mutated genes and regions of gain/loss are colour-coded in the labels (green, loss; red, gain). For example *B2M*, which encodes beta-2-microglobulin, is recurrently mutated and is deleted in two cases.

*IKZF3* and *B2M*) were not previously known targets of point mutation in NHL (Supplementary Table 6 and Supplementary Methods). *FOXO1*, *BCL7A* and *B2M* had hot spots affecting their start codons. The effect of a *FOXO1* start codon mutation, which was observed in three cases, was further studied using a cell line in which the initiating ATG was mutated to TTG. Western blots probed with a *FOXO1* antibody revealed a band with a reduced molecular weight, indicative of a *FOXO1* amino-terminal truncation (Supplementary Figure 2), consistent with use of the next in-frame ATG for translation initiation. A second hot spot in *FOXO1* at T24 was mutated in two cases. T24 is reportedly phosphorylated by AKT subsequent to B-cell receptor (BCR) stimulation<sup>16</sup> inducing *FOXO1* nuclear export.

We analysed the RNA-seq data to determine whether any of the somatic mutations in the 109 recurrently mutated genes showed evidence for allelic imbalance with expression favouring one allele. Out of 380 expressed heterozygous mutant alleles, we observed preferential expression of the mutation for 16.8% (64/380) and preferential expression of the wild type for 27.8% (106/380; Supplementary Table 7). Seven genes showed evidence for significant preferential expression of the mutant allele in at least two cases: *BCL2*, *CARD11*, *CD79B*, *EZH2*, *IRF4*, *MEF2B* and *TP53*; Supplementary Methods. In 27 out of 43 cases with *BCL2* cSNVs, expression favoured the mutant allele, consistent with the previously-described hypothesis that the translocated (and hence, transcriptionally deregulated) allele of *BCL2* is targeted by somatic hypermutation<sup>17</sup>. Examples of mutations at known oncogenic hot spot sites such as F123I in *CARD11* (ref. 18) showed allelic imbalance favouring the mutant allele in some cases. Similarly, we noted expression favouring two novel hot spot mutations in *MEF2B* (Y69 and D83) and two sites in *EZH2* not previously reported as mutated in lymphoma (A682G and A692V).

We sought to distinguish new cancer-related mutations from passenger mutations using the approach proposed previously<sup>19</sup>. We reasoned that this would reveal genes with strong selection signatures, and mutations in such genes would be good candidate cancer drivers. We identified 26 genes with significant evidence for positive selection (false discovery rate = 0.03, Supplementary Methods), with either selective pressure for acquiring non-synonymous point mutations or truncating/nonsense mutations (Supplementary Methods; Table 1 and Supplementary Table 8). Included were known lymphoma oncogenes (*BCL2*, *CD79B* (ref. 9), *CARD11* (ref. 18), *MYD88* (ref. 10) and *EZH2* (ref. 13)), all of which showed signatures indicative of selection for non-synonymous variants.

## Evidence for selection of inactivating changes

We expected tumour suppressor genes to show strong selection for the acquisition of nonsense mutations. In our analysis, the eight most significant genes included seven with strong selective pressure for nonsense mutations, including the known tumour suppressor genes *TP53* and *TNFRSF14* (ref. 20; Table 1). *CREBBP*, recently reported as commonly inactivated in DLBCL<sup>15</sup>, also showed some evidence for acquisition of nonsense mutations and cSNVs (Supplementary Figure 3 and Supplementary Table 9). We also observed enrichment for nonsense mutations in *BCL10*, a positive regulator of NF- $\kappa$ B, in which oncogenic truncated products have been described in lymphomas<sup>21</sup>. The remaining strongly significant genes (*BTG1*, *GNA13*, *SGK1* and *MLL2*) had no reported role in lymphoma. *GNA13* was affected by mutations in 22 cases including multiple nonsense mutations. *GNA13* encodes the alpha subunit of a heterotrimeric G-protein coupled receptor responsible for modulating RhoA activity<sup>22</sup>. Some of the mutated residues negatively affect its function<sup>23,24</sup>, including a T203A mutation, which also showed allelic imbalance favouring the mutant allele (Supplementary Table 7). *GNA13* protein was reduced or absent on western blots in cell lines harbouring either a nonsense mutation, a stop codon deletion, a frame shifting deletion, or changes affecting splice sites (Supplementary Methods and Supplementary Figure 4).

**Table 1 | Overview of cSNVs and confirmed somatic mutations in most frequently mutated genes**

Gene	Cases			Total			Somatic cSNVs (RNA-seq cohort)*	<i>P</i> (raw)	<i>q</i>	NS SP	T SP	Skew (M, WT, both)†
	NS	S	T	NS	S	T						
<i>MLL2</i> ‡	16	8	17	17	8	18	10	<b><math>6.85 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	0.834	14.4	WT
<i>TNFRSF14</i> G‡	7	1	7	8	1	7	11	<b><math>6.85 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	7.52	118	Both
<i>SGK1</i> G‡	18	6	6	37	10	6	9	<b><math>6.85 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	19.5	61.7	—
<i>BCL10</i> ‡	2	0	4	3	0	4	4	<b><math>6.85 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	3.62	112	WT
<i>GNA13</i> G‡	21	1	2	33	1	2	5	<b><math>6.85 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	24.1	25.7	Both
<i>TP53</i> G‡	20	2	1	23	3	1	22	<b><math>6.85 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	15.6	14.1	Both
<i>EZH2</i> G‡	33	0	0	33	0	0	33	<b><math>6.85 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	11.4	0.00	Both
<i>BTG2</i> ‡	12	6	1	14	6	1	2	$6.85 \times 10^{-8}$	$8.50 \times 10^{-7}$	23.9	35.1	—
<i>BCL2</i> G‡	42	45	0	96	105	0	43	<b><math>9.35 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	3.78	0.00	M
<i>BCL6</i> ‡§	11	2	0	12	2	0	2	<b><math>9.35 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	0.175	0.00	M
<i>CIITA</i> ‡§	5	3	0	6	3	0	2	<b><math>9.35 \times 10^{-8}</math></b>	$8.50 \times 10^{-7}$	0.086	0.00	—
<i>FAS</i> ‡	2	0	4	3	0	4	2	$1.52 \times 10^{-7}$	$1.17 \times 10^{-6}$	2.54	66.5	WT
<i>BTG1</i> ‡	11	6	2	11	7	2	10	$1.52 \times 10^{-7}$	$1.17 \times 10^{-6}$	17.5	52.5	Both
<i>MEF2B</i> G‡	20	2	0	20	2	0	10	$2.05 \times 10^{-7}$	$1.47 \times 10^{-6}$	14.2	0.00	M
<i>IRF8</i> ‡	11	5	3	14	5	3	3	$4.55 \times 10^{-7}$	$3.03 \times 10^{-6}$	8.82	28.2	WT
<i>TMEM30A</i> ‡	1	0	4	1	0	4	4	$6.06 \times 10^{-7}$	$3.79 \times 10^{-6}$	0.785	65.0	WT
<i>CD58</i> ‡	2	0	3	2	0	3	2	$2.42 \times 10^{-6}$	$1.43 \times 10^{-5}$	2.29	69.2	—
<i>KLHL6</i> ‡	10	2	2	12	2	2	4	$1.00 \times 10^{-5}$	$5.26 \times 10^{-5}$	5.42	16.4	—
<i>MYD88</i> A‡	13	2	0	14	2	0	9	$1.00 \times 10^{-5}$	$5.26 \times 10^{-5}$	12.4	0.00	WT
<i>CD70</i> ‡	5	0	1	5	0	2	3	$1.70 \times 10^{-5}$	$8.48 \times 10^{-5}$	7.08	44.0	—
<i>CD79B</i> A‡	7	2	1	9	2	1	5	$2.00 \times 10^{-5}$	$9.52 \times 10^{-5}$	10.9	18.3	M
<i>CCND3</i> ‡	7	1	2	7	1	2	6	$2.80 \times 10^{-5}$	$1.27 \times 10^{-4}$	6.55	36.3	WT
<i>CREBBP</i> ‡	20	7	4	24	7	4	9	$1.00 \times 10^{-4}$	$4.35 \times 10^{-4}$	2.72	6.04	Both
<i>HIST1H1C</i> ‡	9	0	0	10	0	0	6	$1.80 \times 10^{-4}$	$7.50 \times 10^{-4}$	11.9	0.00	Both
<i>B2M</i> ‡	7	0	0	7	0	0	4	$3.90 \times 10^{-4}$	$1.56 \times 10^{-3}$	16.6	0.00	WT
<i>ETS1</i> ‡	10	1	0	10	1	0	4	$4.10 \times 10^{-4}$	$1.58 \times 10^{-3}$	5.76	0.00	WT
<i>CARD11</i> ‡	14	3	0	14	3	0	3	$1.90 \times 10^{-3}$	$7.04 \times 10^{-3}$	3.37	0.00	Both
<i>FAT2</i> ‡§	2	1	0	2	1	0	2	$6.30 \times 10^{-3}$	$2.25 \times 10^{-2}$	0.128	0.00	—
<i>IRF4</i> ‡§	9	4	0	26	5	0	5	$7.00 \times 10^{-3}$	$2.41 \times 10^{-2}$	0.569	0.00	Both
<i>FOXO1</i> ‡	8	4	0	10	4	0	4	$7.60 \times 10^{-3}$	$2.53 \times 10^{-2}$	4.02	0.00	—
<i>STAT3</i>	9	0	0	9	0	0	4	$2.19 \times 10^{-2}$	$6.08 \times 10^{-2}$	—	—	Both
<i>RAPGEF1</i>	8	3	0	10	3	0	3	$2.98 \times 10^{-2}$	$7.45 \times 10^{-2}$	—	—	WT
<i>ABCA7</i>	12	3	0	15	3	0	2	$7.76 \times 10^{-2}$	$1.67 \times 10^{-1}$	—	—	WT
<i>RNF213</i>	10	8	0	10	8	0	2	$7.87 \times 10^{-2}$	$1.67 \times 10^{-1}$	—	—	—
<i>MUC16</i>	17	12	0	39	25	0	2	$8.32 \times 10^{-2}$	$1.73 \times 10^{-1}$	—	—	—
<i>HDAC7</i>	8	4	0	8	4	0	2	$8.94 \times 10^{-2}$	$1.82 \times 10^{-1}$	—	—	WT
<i>PRKDC</i>	7	3	0	7	4	0	2	$1.06 \times 10^{-1}$	$2.05 \times 10^{-1}$	—	—	—
<i>SAMD9</i>	9	2	0	9	2	0	2	$1.79 \times 10^{-1}$	$3.01 \times 10^{-1}$	—	—	—
<i>TAF1</i>	10	0	0	10	0	0	2	$3.03 \times 10^{-1}$	$4.74 \times 10^{-1}$	—	—	—
<i>PIM1</i>	20	19	0	33	34	0	11	$3.40 \times 10^{-1}$	$5.23 \times 10^{-1}$	—	—	WT
<i>COL4A2</i>	8	2	0	8	2	0	2	$7.64 \times 10^{-1}$	$8.99 \times 10^{-1}$	—	—	—
<i>EP300</i>	8	7	1	8	7	1	3	$9.54 \times 10^{-1}$	1.00	—	—	WT

Individual cases with non-synonymous (NS), synonymous (S) and truncating (T) mutations and the total number of mutations of each class are shown separately because some genes contained multiple mutations in the same case. The *P* values indicated in bold are the upper limit on the *P* value for that gene determined with the approach described in ref. 19 (see Supplementary Methods), *q* is the Benjamini-corrected *q* value, and NS SP and T SP refer to selective pressure estimates from this model for the acquisition of non-synonymous or truncating mutations, respectively. Genes with a superscript of either A or G were found to have mutations significantly enriched in ABC or GCB cases, respectively ( $P < 0.05$ , Fisher's exact test).

\* Additional somatic mutations identified in larger cohorts and insertion/deletion mutations are not included in this total.

† 'Both' indicates that we observed separate cases in which skewed expression was seen but where this skew was not consistent for the mutant or wild-type allele.

‡ Genes significant at a false discovery rate of 0.03. SNVs in *BCL2* and previously confirmed hot spot mutations in *EZH2* and *CD79B* are probably somatic in these samples based on published observations of others.

§ Selective pressure estimates are both  $< 1$  indicating purifying selection rather than positive selection acting on this gene.

*SGK1* encodes a phosphatidylinositol-3-OH kinase (PI(3)K)-regulated kinase with functions including regulation of FOXO transcription factors<sup>25</sup>, regulation of NF-κB by phosphorylating IκB kinase<sup>26</sup>, and negative regulation of NOTCH signalling<sup>27</sup>. *SGK1* also resides within a region of chromosome 6 commonly deleted in DLBCL (Fig. 1)<sup>5</sup>. The mechanism by which *SGK1* and *GNA13* inactivation may contribute to lymphoma is unclear, but the strong degree of apparent selection towards their inactivation and their overall high mutation frequency (each mutated in 18 of 106 DLBCL cases) suggests that their loss contributes to B-cell NHL. Certain genes are known to be mutated more commonly in GCB DLBCLs (for example, *TP53* (ref. 28) and *EZH2* (ref. 13)). Here, both *SGK1* and *GNA13* mutations were found only in GCB cases ( $P = 1.93 \times 10^{-3}$  and  $2.28 \times 10^{-4}$ , Fisher's exact test;  $n = 15$  and 18, respectively) (Fig. 2). Two additional genes (*MEF2B* and *TNFRSF14*) with no previously described role in DLBCL showed a similar restriction to GCB cases (Fig. 2).

### Inactivating *MLL2* mutations

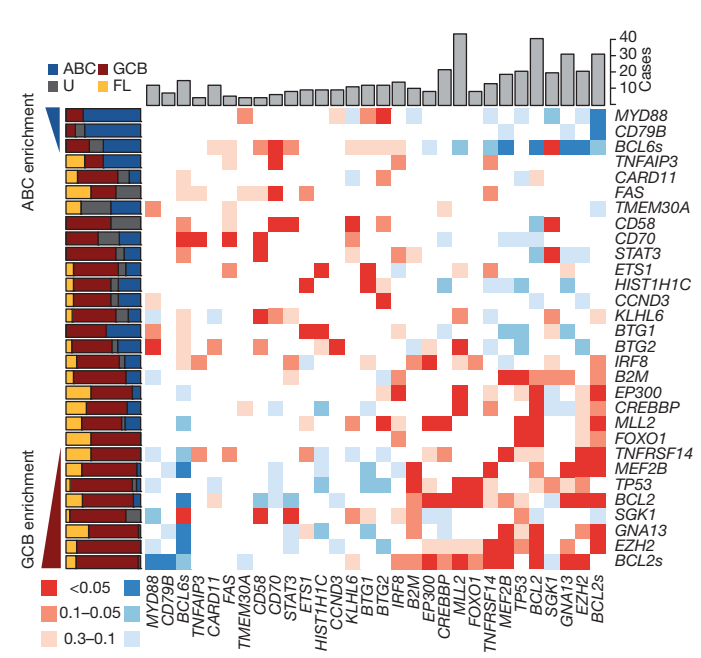
*MLL2* showed the most significant evidence for selection and the largest number of nonsense SNVs. Our RNA-seq analysis indicated that 26.0% (33/127) of cases carried at least one *MLL2* cSNV. To

address the possibility that variable RNA-seq coverage of *MLL2* failed to capture some mutations, we PCR-amplified the entire *MLL2* locus (~36 kilobases) in 89 cases (35 primary FLs, 17 DLBCL cell lines, and 37 DLBCLs). Of these cases 58 were among the RNA-seq cohort. Illumina amplicon re-sequencing (Supplementary Methods) revealed 78 mutations, confirming the RNA-seq mutations in the overlapping cases and identifying 33 additional mutations. We confirmed the somatic status of 46 variants using Sanger sequencing (Supplementary Table 10), and showed that 20 of the 33 additional mutations were insertions or deletions (indels). Three SNVs at splice sites were also detected, as were 10 new cSNVs that had not been detected by RNA-seq.

The somatic mutations were distributed across *MLL2* (Fig. 3a). Of these, 37% ( $n = 29/78$ ) were nonsense mutations, 46% ( $n = 36/78$ ) were indels that altered the reading frame, 8% ( $n = 6/78$ ) were point mutations at splice sites and 9% ( $n = 7/78$ ) were non-synonymous amino acid substitutions (Table 2). Four of the somatic splice site mutations had effects on *MLL2* transcript length and structure. For example, two heterozygous splice site mutations resulted in the use of a novel splice donor site and an intron retention event.

Approximately half of the NHL cases we sequenced had two *MLL2* mutations (Supplementary Table 10). We used bacterial artificial





**Figure 2 | Overview of mutations and potential cooperative interactions in NHL.** This heat map displays possible trends towards co-occurrence (red) and mutual exclusion (blue) of somatic mutations and structural rearrangements. Colours were assigned by taking the minimum value of a left- and right-tailed Fisher's exact test. To capture trends a *P*-value threshold of 0.3 was used, with the darkest shade of the colour indicating those meeting statistical significance ( $P \leq 0.05$ ). The relative frequency of mutations in ABC (blue), GCB (red), unclassifiable (black) DLBCLs and FL (yellow) cases is shown on the left. Genes were arranged with those having significant ( $P < 0.05$ , Fisher's exact test) enrichment for mutations in ABC cases (blue triangle) towards the top (and left) and those with significant enrichment for mutations in GCB cases (red triangle) towards the bottom (and right). The total number of cases in which each gene contained either cSNVs or confirmed somatic mutations is shown at the top. The cluster of blue squares (upper-right) results from the mutual exclusion of the ABC-enriched mutations (for example, *MYD88*, *CD79B*) from the GCB-enriched mutations (for example, *EZH2*, *GNA13*). Presence of structural rearrangements involving the two oncogenes *BCL6* and *BCL2* (indicated as *BCL6s* and *BCL2s*) was determined with FISH techniques using break-apart probes (Supplementary Methods).

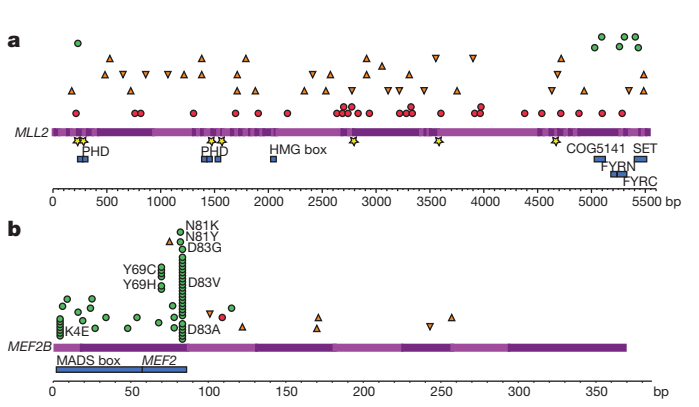
chromosome (BAC) clone sequencing in eight FL cases to show that in all eight cases the mutations were in *trans*, affecting both *MLL2* alleles. This observation is consistent with the notion that there is a complete, or near-complete, loss of *MLL2* in the tumour cells of such patients.

With the exception of two primary FL cases and two DLBCL cell lines (Pfeiffer and SU-DHL-9), the majority of *MLL2* mutations seemed to be heterozygous. Analysis of Affymetrix 500k SNP array data from two FL cases with apparent homozygous mutations revealed that both tumours showed copy number neutral loss of heterozygosity (LOH) for the region of chromosome 12 containing *MLL2* (Supplementary Methods). Thus, in addition to bi-allelic mutation, LOH is a second, albeit less common mechanism by which *MLL2* function is lost.

*MLL2* was the most frequently mutated gene in FL, and among the most frequently mutated genes in DLBCL (Fig. 2). We confirmed *MLL2* mutations in 31 of 35 FL patients (89%), in 12 of 37 DLBCL patients (32%), in 10 of 17 DLBCL cell lines (59%) and in none of the eight normal centroblast samples we sequenced. Our analysis predicted that the majority of the somatic mutations observed in *MLL2* were inactivating (91% disrupted the reading frame or were truncating point mutations), indicating to us that *MLL2* is a tumour suppressor of significance in NHL.

### Recurrent point mutations in *MEF2B*

Our selective pressure analysis also revealed genes with stronger pressure for acquisition of amino acid substitutions than for nonsense



**Figure 3 | Summary and effect of somatic mutations affecting *MLL2* and *MEF2B*.** **a**, Re-sequencing the *MLL2* locus in 89 samples revealed mainly nonsense (red circles) and frameshift-inducing indel mutations (orange triangles; inverted triangles for insertions and upright triangles for deletions). A smaller number of non-synonymous somatic mutations (green circles) and point mutations or deletions affecting splice sites (yellow stars) were also observed. All of the non-synonymous point mutations affected a residue within either the catalytic SET domain, the FYRC domain (FY-rich carboxy-terminal domain) or PHD zinc finger domains. The effect of these splice-site mutations on *MLL2* splicing was also explored (Supplementary Figure 7). **b**, The cSNVs and somatic mutations found in *MEF2B* in all FL and DLBCL cases sequenced are shown with the same symbols. Only the amino acids with variants in at least two patients are labelled. cSNVs were most prevalent in the first two protein-coding exons of *MEF2B* (exons 2 and 3). The crystal structure of MEF2 bound to EP300 supports the idea that two of the mutated sites (L67 and Y69) are important in the interaction between these proteins (Supplementary Figure 8 and Supplementary Discussion)<sup>50</sup>.

mutations. One such gene was *MEF2B*, which had not previously been linked to lymphoma. We found that 20 (15.7%) cases had *MEF2B* cSNVs and 4 (3.1%) cases had *MEF2C* cSNVs. All cSNVs detected by RNA-seq affected either the MADS box or MEF2 domains. To determine the frequency and scope of *MEF2B* mutations, we Sanger-sequenced exons 2 and 3 in 261 primary FL samples; 259 DLBCL primary tumours; 17 cell lines; 35 cases of assorted NHL (IBL, composite FL and PBMCL); and eight non-malignant centroblast samples. We also used a capture strategy (Supplementary Methods) to sequence the entire *MEF2B* coding region in the 261 FL samples, revealing six additional variants outside exons 2 and 3. We thus identified 69 cases (34 DLBCL, 12.67%; and 35 FL, 15.33%) with *MEF2B* cSNVs or indels, failing to observe novel variants in other NHL and non-malignant samples. Of the variants 55 (80%) affected residues within the MADS box and MEF2 domains encoded by exons 2 and 3 (Supplementary Table 11; Fig. 3b). Each patient generally had a single *MEF2B* variant and we observed relatively few (eight in total, 10.7%) truncation-inducing SNVs or indels. Non-synonymous SNVs were by far the most common type of change observed, with 59.4% of detected variants affecting K4, Y69, N81 or D83. In 12 cases *MEF2B* mutations were shown to be somatic, including representative mutations at each of K4, Y69, N81 and D83 (Supplementary Table 12). We did not detect mutations in ABC cases, indicating that somatic mutations in *MEF2B* have a role unique to the development of GCB DLBCL and FL (Fig. 2).

**Table 2 | Summary of types of *MLL2* somatic mutations**

Sample Type	FL	DLBCL	DLBCL cell-line	Centroblast
Truncation	18	4	7	0
Indel with frameshift	22	8	6	0
Splice site	4	2	0	0
SNV	3	2	2	0
Any mutation/ number of cases	31/35	12/37	10/17	0/8
Percentage	89	32	59	0

## Discussion

In our study of genome, transcriptome and exome sequences from 127 B-cell NHL cases, we identified 109 genes with clear evidence of somatic mutation in multiple individuals. Significant selection seems to act on at least 26 of these for the acquisition of either nonsense or missense mutations. To the best of our knowledge, the majority of these genes had not previously been associated with any cancer type. We observed an enrichment of somatic mutations affecting genes involved in transcriptional regulation and, more specifically, chromatin modification.

*MLL2* emerged from our analysis as a major tumour suppressor locus in NHL. It is one of six human H3K4-specific methyltransferases<sup>29</sup>, all of which share homology with the *Drosophila* trithorax gene. Trimethylated H3K4 (H3K4me3) is an epigenetic mark associated with the promoters of actively transcribed genes. By laying down this mark, MLLs are responsible for the transcriptional regulation of developmental genes including the homeobox (*Hox*) gene family<sup>30</sup> which collectively control segment specificity and cell fate in the developing embryo<sup>31,32</sup>. Each MLL family member is thought to target different subsets of *Hox* genes<sup>33</sup> and in addition, *MLL2* is known to regulate the transcription of a diverse set of genes<sup>34</sup>. Recently, *MLL2* mutations were reported in a small-cell lung cancer cell line<sup>35</sup> and in renal carcinoma<sup>36</sup>, but the frequency of nonsense mutations affecting *MLL2* in these cancers was not established in these reports. Inactivating mutations were reported recently in *MLL2* or *MLL3* in 16% of medulloblastoma patients<sup>37</sup>, further implicating *MLL2* as a cancer gene.

Our data link *MLL2* somatic mutations to B-cell NHL. The reported mutations are likely to be inactivating and in eight of the cases with multiple mutations, we confirmed that both alleles were affected, presumably resulting in essentially complete loss of *MLL2* function. The high prevalence of *MLL2* mutations in FL (89%) equals the frequency of the t(14;18)(q32;q21) translocation, which is considered the most prevalent genetic abnormality in FL<sup>3</sup>. In DLBCL tumour samples and cell lines, *MLL2* mutation frequencies were 32% and 59%, respectively, also exceeding the prevalence of the most frequent cytogenetic abnormalities, such as the various translocations involving 3q27, which occur in 25–30% of DLBCLs and are enriched in ABC cases<sup>38</sup>. Importantly, we found *MLL2* mutated in both DLBCL subtypes (Fig. 2). Our analyses thus indicate that *MLL2* acts as a central tumour suppressor in FL and both DLBCL subtypes.

The *MEF2* gene family encodes four related transcription factors that recruit histone-modifying enzymes including histone deacetylases (HDACs) and HATs in a calcium-regulated manner. Although truncating variants were detected in our analysis of *MEF2* gene family members, our analysis suggests that, in contrast to *MLL2*, *MEF2* family members tend to selectively acquire non-synonymous amino acid substitutions. In the case of *MEF2B*, 59.4% of all the cSNVs were found at four sites within the protein (K4, Y69, N81 and D83), and all four of these sites were confirmed to be targets of somatic mutation. D83 is affected in 39% of the *MEF2B* alterations, resulting in replacement of the charged aspartate with any of alanine, glycine or valine. Although we cannot yet predict the consequences of these substitutions on protein function, it seems likely that their effect would have an impact on the ability of *MEF2B* to facilitate gene expression and thus have a role in promoting the malignant transformation of germinal centre B cells to lymphoma (Supplementary Discussion).

*MEF2B* mutations can be linked to *CREBBP* and *EP300* mutations, and to recurrent Y641 mutations in *EZH2* (ref. 13). One target of *CREBBP*/*EP300* HAT activity is H3K27, which is methylated by *EZH2* to repress transcription. There is evidence that the action of *EZH2* antagonizes that of *CREBBP*/*EP300* (ref. 39). One function of *MEF2* is to recruit either HDACs or *CREBBP*/*EP300* to target genes<sup>40</sup>, and it has been suggested that HDACs compete with *CREBBP*/*EP300* for the same binding site on *MEF2* (ref. 41). Under normal  $\text{Ca}^{2+}$  levels, *MEF2* is bound by type IIa HDACs, which maintain the tails

of histone proteins in a deacetylated repressive chromatin state<sup>42</sup>. Increased cytoplasmic  $\text{Ca}^{2+}$  levels induce the nuclear export of HDACs, enabling the recruitment of HATs such as *CREBBP*/*EP300*, facilitating transcription at *MEF2* target genes. Mutation of *CREBBP*, *EP300* or *MEF2B* may have an impact on the expression of *MEF2* target genes owing to reduced acetylation of nucleosomes near these genes (Supplementary Figure 5; Supplementary Discussion). In light of the recent finding that heterozygous *EZH2* Y641 mutations enhance overall H3K27 trimethylation activity of PRC2 (refs 43, 44), it is possible that mutation of both *MLL2* and *EZH2* could cooperate in reducing the expression of some of the same target genes. Our data indicate that (1) post-transcriptional modification of histones is of key importance in germinal centre B cells and (2) deregulated histone modification due to these mutations is likely to result in reduced acetylation and enhanced methylation, and acts as a core driver event in the development of NHL (Supplementary Figure 5).

## METHODS SUMMARY

All samples analysed contained at least 50% tumour cells. Genomes, exomes and transcriptomes were sequenced using a combination of Illumina GAIIx and HiSeq 2000 instruments to read lengths of between 36 and 100 nucleotides. Exome capture was performed using the Agilent SureSelect Target Enrichment System Protocol (Version 1.0, September 2009). Alignment was accomplished using BWA<sup>45</sup> and variants were identified using SNVMix<sup>46</sup>. Variants were manually reviewed in IGV and were confirmed (where applicable) by PCR followed by either Sanger sequencing or Illumina re-sequencing. Structural rearrangements in genomes and transcriptomes were identified using ABySS<sup>47</sup>. Gene expression values used for subtype assignment were calculated as reads per kilobase gene model per million mapped reads (RPKM) values<sup>48</sup> and subtypes were assigned using an adaptation of the method developed for data from Affymetrix expression arrays<sup>49</sup> trained with samples previously classified by this standard approach.

Received 13 November 2010; accepted 7 July 2011.

Published online 27 July 2011.

- Anderson, J. R., Armitage, J. O., Weisenburger, D. D., Non-Hodgkin's Lymphoma Classification Project. Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. *Ann. Oncol.* **9**, 717–720 (1998).
- Lenz, G. & Staudt, L. M. Aggressive lymphomas. *N. Engl. J. Med.* **362**, 1417–1429 (2010).
- Horsman, D. E. *et al.* Follicular lymphoma lacking the t(14;18)(q32;q21): identification of two disease subtypes. *Br. J. Haematol.* **120**, 424–433 (2003).
- Iqbal, J. *et al.* *BCL2* translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma. *Am. J. Pathol.* **165**, 159–166 (2004).
- Lenz, G. *et al.* Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc. Natl Acad. Sci. USA* **105**, 13520–13525 (2008).
- Pasqualucci, L. *et al.* Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J. Exp. Med.* **203**, 311–317 (2006).
- Kato, M. *et al.* Frequent inactivation of A20 in B-cell lymphomas. *Nature* **459**, 712–716 (2009).
- Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF- $\kappa$ B in diffuse large B-cell lymphoma. *Nature* **459**, 717–721 (2009).
- Davis, R. E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463**, 88–92 (2010).
- Ngo, V. N. *et al.* Oncogenically active MYD88 mutations in human lymphoma. *Nature* **470**, 115–119 (2011).
- Mardis, E. R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* **361**, 1058–1066 (2009).
- Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–813 (2009).
- Morin, R. D. *et al.* Somatic mutations altering *EZH2* (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature Genet.* **42**, 181–185 (2010).
- Futreal, P. A. *et al.* A census of human cancer genes. *Nature Rev. Cancer* **4**, 177–183 (2004).
- Pasqualucci, L. *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* **471**, 189–195 (2011).
- Yusuf, I., Zhu, X., Kharas, M. G., Chen, J. & Fruman, D. A. Optimal B-cell proliferation requires phosphoinositide 3-kinase-dependent inactivation of FOXO transcription factors. *Blood* **104**, 784–787 (2004).
- Saito, M. *et al.* BCL6 suppression of *BCL2* via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc. Natl Acad. Sci. USA* **106**, 11294–11299 (2009).
- Lenz, G. *et al.* Oncogenic *CARD11* mutations in human diffuse large B cell lymphoma. *Science* **319**, 1676–1679 (2008).

19. Greenman, C., Wooster, R., Futreal, P. A., Stratton, M. R. & Easton, D. F. Statistical analysis of pathogenicity of somatic mutations in cancer. *Genetics* **173**, 2187–2198 (2006).
20. Cheung, K. J. *et al.* Acquired *TNFRSF14* mutations in follicular lymphoma are associated with worse prognosis. *Cancer Res.* **70**, 9166–9174 (2010).
21. Du, M. Q. *et al.* *BCL10* gene mutation in lymphoma. *Blood* **95**, 3885–3890 (2000).
22. Kreutz, B., Hajicek, N., Yau, D. M., Nakamura, S. & Kozasa, T. Distinct regions of G $\alpha$ 13 participate in its regulatory interactions with RGS homology domain-containing RhoGEFs. *Cell. Signal.* **19**, 1681–1689 (2007).
23. Bhattacharyya, R. & Wedegaertner, P. G $\alpha$ 13 requires palmitoylation for plasma membrane localization, Rho-dependent signaling, and promotion of p115-RhoGEF membrane binding. *J. Biol. Chem.* **275**, 14992–14999 (2000).
24. Manganello, J. M., Huang, J., Kozasa, T., Voyno-Yasenetskaya, T. A. & Le Breton, G. C. Protein kinase A-mediated phosphorylation of the G $\alpha$ 13 switch I region alters the G $\alpha$ 13-G $\beta$ 13-G protein-coupled receptor complex and inhibits Rho activation. *J. Biol. Chem.* **278**, 124–130 (2003).
25. Brunet, A. *et al.* Protein kinase SGK mediates survival signals by phosphorylating the forkhead transcription factor FKHL1 (FOXO3a). *Mol. Cell. Biol.* **21**, 952–965 (2001).
26. Tai, D. J. C., Su, C.-C., Ma, Y.-L. & Lee, E. H. Y. SGK1 phosphorylation of I $\kappa$ B kinase  $\alpha$  and p300 Up-regulates NF- $\kappa$ B activity and increases N-methyl-D-aspartate receptor NR2A and NR2B expression. *J. Biol. Chem.* **284**, 4073–4089 (2009).
27. Mo, J. *et al.* Serum- and glucocorticoid-inducible kinase 1 (SGK1) controls Notch1 signaling by downregulation of protein stability through Fbw7 ubiquitin ligase. *J. Cell Sci.* **124**, 100–112 (2011).
28. Young, K. H. *et al.* Structural profiles of *TP53* gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood* **112**, 3088–3098 (2008).
29. Shilatifard, A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr. Opin. Cell Biol.* **20**, 341–348 (2008).
30. Milne, T. *et al.* MLL targets SET domain methyltransferase activity to *Hox* gene promoters. *Mol. Cell* **10**, 1107–1117 (2002).
31. Krumlauf, R. *Hox* genes in vertebrate development. *Cell* **78**, 191–201 (1994).
32. Canaan, E. *et al.* ALL-1/MLL1, a homologue of *Drosophila* TRITHORAX, modifies chromatin and is directly involved in infant acute leukaemia. *Br. J. Cancer* **90**, 756–760 (2004).
33. Wang, P. *et al.* Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol. Cell. Biol.* **29**, 6074–6085 (2009).
34. Issaeva, I. *et al.* Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol. Cell. Biol.* **27**, 1889–1903 (2007).
35. Pleasance, E. D. *et al.* A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463**, 184–190 (2010).
36. Dalglish, G. L. *et al.* Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360–363 (2010).
37. Parsons, D. W. *et al.* The genetic landscape of the childhood cancer medulloblastoma. *Science* **331**, 435–439 (2011).
38. Iqbal, J. *et al.* Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B-cell lymphoma. *Leukemia* **21**, 2332–2343 (2007).
39. Pasini, D. *et al.* Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res.* (2010).
40. Giordano, A. & Avantiaggiati, M. p300 and CBP: partners for life and death. *J. Cell. Physiol.* **181**, 218–230 (1999).
41. Han, A., He, J., Wu, Y., Liu, J. O. & Chen, L. Mechanism of recruitment of class II histone deacetylases by myocyte enhancer factor-2. *J. Mol. Biol.* **345**, 91–102 (2005).
42. Youn, H. & Liu, J. Cabin1 represses MEF2-dependent Nur77 expression and T cell apoptosis by controlling association of histone deacetylases and acetylases with MEF2. *Immunity* **13**, 85–94 (2000).
43. Yap, D. B. *et al.* Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood* **117**, 2451–2459 (2011).
44. Sneringer, C. J. *et al.* Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc. Natl Acad. Sci. USA* **107**, 20980–20985 (2010).
45. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
46. Goya, R. *et al.* SNVMix: predicting single nucleotide variants from next-generation sequencing of tumors. *Bioinformatics* **26**, 730–736 (2010).
47. Robertson, G. *et al.* De novo assembly and analysis of RNA-seq data. *Nature Methods* **7**, 909–912 (2010).
48. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621–628 (2008).
49. Wright, G. *et al.* A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc. Natl Acad. Sci. USA* **100**, 9991–9996 (2003).
50. He, J. *et al.* Structure of p300 bound to MEF2 on DNA reveals a mechanism of enhanceosome assembly. *Nucleic Acids Res.* (2011).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This study was funded in part by funding from the National Cancer Institute Office of Cancer Genomics (Contract No. HHSN261200800001E), the Terry Fox Foundation (grant 019001, Biology of Cancer: Insights from Genomic Analyses of Lymphoid Neoplasms) and Genome Canada/Genome British Columbia Grant Competition III (Project Title: High Resolution Analysis of Follicular Lymphoma Genomes) to J.M.C., R.D.G. and M.A.M. We acknowledge support from NIH grants P50CA130805-01 “SPORE in Lymphoma, Tissue Resource Core (PI Fisher)” and 1U01CA114778 “Molecular Signatures to Improve Diagnosis and Outcome in Lymphoma (PI Chan)”. A.J.M. is a Career Development Program Fellow of the Leukemia and Lymphoma Society. N.A.J. was a research fellow of the Terry Fox Foundation (award NCIC 019005) and the Michael Smith Foundation for Health Research (ST-PDF-01793). M.A.M. is a Terry Fox Young Investigator and a Michael Smith Senior Research Scholar. R.D.M. is a Vanier Scholar (CIHR) and holds a MSFHR senior graduate studentship. M.M.-L. acknowledges support from a Postdoctoral Fellowship from the Spanish Ministry of Education, under the “Programa Nacional de Movilidad de Recursos Humanos del Plan Nacional de I-D+i 2008-2011”. D.W.S. was supported by the Terry Fox Foundation Strategic Health Research Training Program in Cancer Research at Canadian Institutes of Health Research (Grant No. TGT-53912). J.J.S. acknowledges funding from The Canadian Cancer Society and the Canadian Institutes of Health Research. R.G. is supported by a UBC Four Year Fellowship. I.M.M. acknowledges the Canadian Foundation for Innovation for a Leaders Opportunity Fund. The laboratory work for this study was undertaken at the Genome Sciences Centre, British Columbia Cancer Research Centre and the Centre for Translational and Applied Genomics, a program of the Provincial Health Services Authority Laboratories. The authors would like to thank C. Greenman for supplying his software and also acknowledge D. Gerhard and S. Aparicio for discussions and guidance. Special thanks to C. Suragh, R. Roscoe, A. Troussard and A. Drobnies for expert project management assistance, and to the Library Construction, Sequencing and Bioinformatics teams at the Genome Sciences Centre. The content of this publication does not necessarily reflect the views of policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

**Author Contributions** M.A.M., R.D.G., D.E.H., M.H. and J.M.C. conceived of the study and led the design of the experiments. R.D.M. performed the analysis of sequence data, identified mutations and, with M.M.-L., A.J.M. and M.A.M., produced figures and wrote the manuscript. M.M.-L., A.J.M., D.L.T., S.Chan, S.Chittarajan, D.S., H.M., J.S., M.M., T.Z., A.D., K.T., Y.B., M.R.F., J.T.-W. and T.M.S. designed and performed experiments to amplify, discover and validate mutations. R.G., M.G. and I.M.M. contributed to analyses and reviewed the manuscript. N.A.J., M.B., B.W. and B.M. prepared the samples, performed sample sorting and COO analysis and contributed to the text. A.B.-W. and J.J.S. collected and prepared constitutional DNA samples. K.L.M., R.C., S.L., M.F. and S.J. generated *de novo* assemblies and identified mutations. M.K., S.R., M.G., O.Y. and E.Y.Z. wrote software and contributed to figures. R.D.C. performed copy number analysis and produced a figure and S.B.-N. performed confirmatory FISH experiments. Y.Z. and A.T. produced the sequencing libraries. I.B., R.H., S.J.M.J., R.M., J.S. and M.H. contributed to the development of experimental and analytical protocols. L.R. provided materials and reviewed the manuscript.

**Author Information** The SRA accession number for the submission of the data not included in previous publications is SRP001599, which is linked to the dbGAP study accession phs000235.v2.p1. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at [www.nature.com/nature](http://www.nature.com/nature). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to M.A.M. ([mmarra@bcgsc.ca](mailto:mmarra@bcgsc.ca)).



# Neocortical excitation/inhibition balance in information processing and social dysfunction

Ofer Yizhar<sup>1,2\*</sup>, Lief E. Fenno<sup>1,3\*</sup>, Matthias Prigge<sup>4</sup>, Franziska Schneider<sup>4</sup>, Thomas J. Davidson<sup>1</sup>, Daniel J. O'Shea<sup>1,3</sup>, Vikaas S. Sohal<sup>1,5</sup>, Inbal Goshen<sup>1</sup>, Joel Finkelstein<sup>1</sup>, Jeanne T. Paz<sup>6</sup>, Katja Stehfest<sup>4</sup>, Roman Fudim<sup>4</sup>, Charu Ramakrishnan<sup>1</sup>, John R. Huguenard<sup>6</sup>, Peter Hegemann<sup>4</sup> & Karl Deisseroth<sup>1,7,8,9</sup>

Severe behavioural deficits in psychiatric diseases such as autism and schizophrenia have been hypothesized to arise from elevations in the cellular balance of excitation and inhibition (E/I balance) within neural microcircuitry. This hypothesis could unify diverse streams of pathophysiological and genetic evidence, but has not been susceptible to direct testing. Here we design and use several novel optogenetic tools to causally investigate the cellular E/I balance hypothesis in freely moving mammals, and explore the associated circuit physiology. Elevation, but not reduction, of cellular E/I balance within the mouse medial prefrontal cortex was found to elicit a profound impairment in cellular information processing, associated with specific behavioural impairments and increased high-frequency power in the 30–80 Hz range, which have both been observed in clinical conditions in humans. Consistent with the E/I balance hypothesis, compensatory elevation of inhibitory cell excitability partially rescued social deficits caused by E/I balance elevation. These results provide support for the elevated cellular E/I balance hypothesis of severe neuropsychiatric disease-related symptoms.

The neurophysiological substrates of most psychiatric disorders are poorly understood, despite accumulating information on genetic factors associated with complex behavioural phenotypes such as those observed in autism and schizophrenia<sup>1–4</sup>. One emerging principle is that a range of seemingly unrelated genetic abnormalities can give rise to the same class of psychiatric phenotype (such as social behaviour dysfunction)<sup>5</sup>, highlighting the need to identify simplifying circuit-level concepts<sup>6–8</sup> that could unify diverse genetic factors under a common pathophysiological principle. One such hypothesis is that elevation in the ratio of cortical cellular excitation to inhibition (cellular E/I balance), for example through increased activity in excitatory neurons or reduction in inhibitory neuron function, could give rise to the social and cognitive deficits observed in diseases such as autism<sup>6,7,9,10</sup>. This hypothesis would unify diverse streams of pathophysiological evidence<sup>11–15</sup>, including the suggestion that many autism-related genes are linked to gain-of-function phenotypes in ion channels and synaptic proteins<sup>16</sup>, that the autistic state is associated with hyperactivity in frontal brain regions<sup>17–19</sup> with elevated high-frequency cortical oscillations<sup>20,21</sup> and that ~30% of autistic patients show clinically apparent seizures<sup>22,23</sup>. However, it is not clear if such an imbalance would impart its disease-relevant effects by influencing development<sup>9,10,24</sup> or real-time function; moreover, the hypothesis has been neither directly testable nor universally accepted<sup>25–27</sup>, although much pioneering work<sup>8–30</sup> may have a bearing on this question.

Although the motivations for social interaction may vary among species, a common theme (independent of motivation) is that social interactions are challenging owing to high levels of uncertainty, and can require rapid interpretation of complex social stimuli to inform

the moment-to-moment generation, refinement and selection of models for the actions of other individuals. Social behaviour assays for experimental mammals such as mice are available, as well as optogenetic methods for direct causal testing of hypotheses regarding the role of activity in specified neuronal populations<sup>31,32</sup>. Using optogenetics, selectively favouring excitation of one population over another (for example with a bistable step-function opsin (SFO), which can depolarize neurons for prolonged periods<sup>33,34</sup>) could address the E/I balance hypothesis, but membrane depolarizations induced by known SFOs<sup>33,35</sup> are not sufficiently stable after a single light flash over the period required for complex behavioural testing. Nor has it been possible thus far to independently excite multiple cell groups within the same volume of tissue, important for testing aspects of the balance hypothesis. Potential strategies exist to manipulate independent neuronal populations in the context of known potent blue-light-activated (~460 nm peak) excitatory tools<sup>36</sup>, but currently available redshifted channelrhodopsins are not suitable for combinatorial control because redshifts of less than 50 nm<sup>37,38</sup> do not translate into separable channels of control *in vivo*, and the most redshifted channelrhodopsin (>70 nm; VChR1) suffers from low expression levels and weak photocurrents<sup>39</sup>.

To enable combinatorial interrogation of E/I balance in experimental animals, we therefore first created several novel optogenetic tools: a series of chimaeric redshifted opsins comprised of sequences from ChR1<sup>40</sup> and VChR1<sup>39</sup>, and a stable SFO operative over the 30 min timescale. We then directly tested the E/I balance hypothesis by optogenetically elevating cellular E/I balance in medial prefrontal

<sup>1</sup>Department of Bioengineering, W083 Clark Center, 318 Campus Drive West, Stanford University, Stanford, California USA. <sup>2</sup>Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel. <sup>3</sup>Neuroscience Program, W083 Clark Center, 318 Campus Drive West, Stanford University, Stanford, California 94305, USA. <sup>4</sup>Institute of Biology, Experimental Biophysics, Humboldt-Universität, Invalidenstraße 42, D-10115 Berlin, Germany. <sup>5</sup>Department of Psychiatry, University of California, San Francisco, San Francisco, California 94143, USA. <sup>6</sup>Department of Neurology and Neurological Science, W083 Clark Center, 318 Campus Drive West, Stanford University, Stanford, California 94305, USA. <sup>7</sup>Howard Hughes Medical Institute, W083 Clark Center, 318 Campus Drive West, Stanford University, Stanford, California 94305, USA. <sup>8</sup>Department of Psychiatry and Behavioral Sciences, W083 Clark Center, 318 Campus Drive West, Stanford University, Stanford, California 94305, USA. <sup>9</sup>CNC Program, W083 Clark Center, 318 Campus Drive West, Stanford University, Stanford, California 94305, USA.

\*These authors contributed equally to this work.

cortex (mPFC) during behaviours relevant for social function and learning in freely moving mice.

## Development of a stabilized SFO

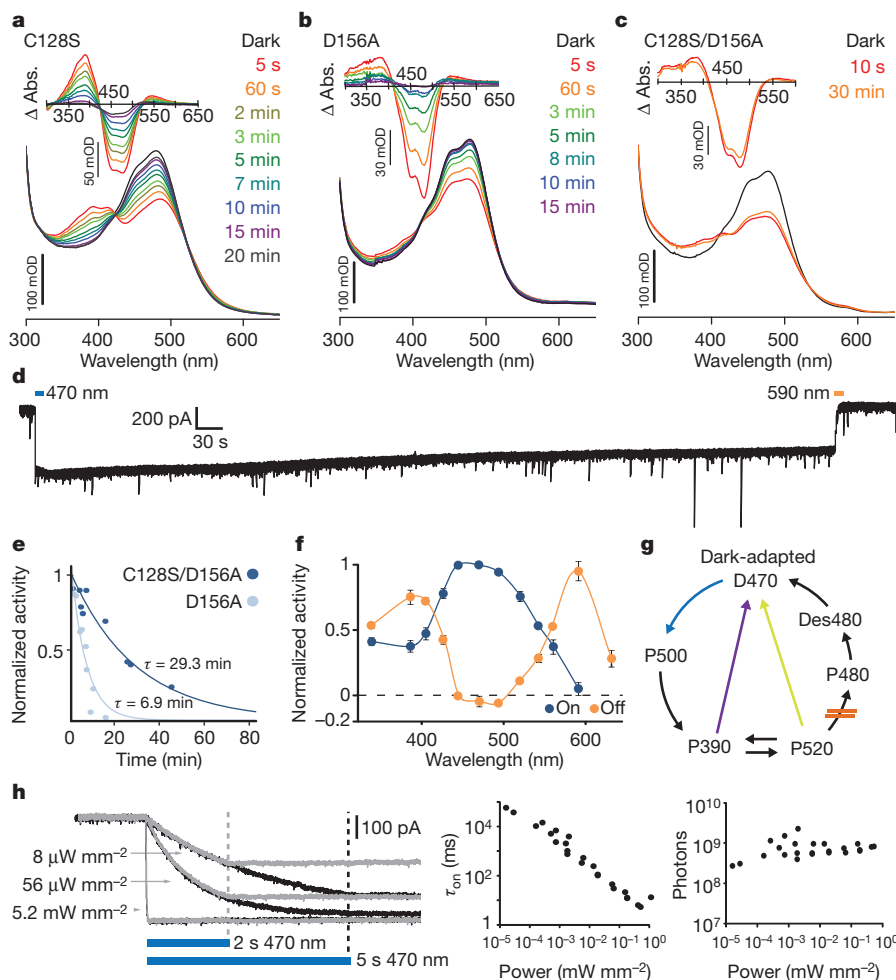
We first sought to generate an SFO suitably stable for mammalian behaviour. We hypothesized that combining previously characterized SFO mutations (C128S<sup>33</sup> and D156A<sup>35</sup>), both shown to confer minute-scale deactivation kinetics, would synergistically stabilize the open channel state. We generated and purified from *Pichia pastoris* the ChR2 mutants C128S, D156A and C128S/D156A (Supplementary Fig. 1a) to measure intrinsic open-state stability in the absence of potentially confounding cellular properties. Time-resolved spectroscopy to measure photocycle progression showed expected rapid absorbance changes with light delivery that largely recovered within several minutes for single mutants C128S (Fig. 1a) and D156A (Fig. 1b). In contrast, the double mutant C128S/D156A showed remarkable stability of the activated state, with little detectable return to the dark state even after 30 min (Fig. 1c and Supplementary Fig. 1b). We next performed whole-cell recordings from neurons expressing this construct (Supplementary Fig. 1c). Consistent with spectroscopic observations, neurons expressing ChR2(C128S/D156A) gave rise to far more stable photocurrents than those induced by either single mutant<sup>33,35</sup> (Fig. 1d). A spontaneous deactivation time constant of 29 min was observed for ChR2(C128S/D156A) ( $r^2 = 0.91$ ), 4.2-fold longer than D156A (6.9 min,  $r^2 = 0.83$ ) (Fig. 1e). Peak photocurrents in neurons expressing ChR2(C128S/D156A) or ChR2(D156A) were comparable ( $231.08 \pm 31.19$  pA;  $n = 9$  cells and  $320.96 \pm 78.26$  pA;  $n = 7$  cells, respectively;  $P = 0.26$ ). The double mutant could be deactivated with either 590 or 405 nm light (Fig. 1f, g and Supplementary

Fig. 1d, e). We refer to the double-mutant gene and protein as the stable SFO (SSFO).

As slow channelrhodopsins enable transduced cells to act as photon integrators across time<sup>33</sup>, we hypothesized that the SSFO would enable photoactivation with extremely low light power. Indeed, intensities as low as  $8 \mu\text{W mm}^{-2}$  generated hundreds of picoamperes of whole-cell photocurrent in neurons expressing SSFO (Fig. 1h). Photocurrent amplitude increased with monoexponential kinetics during illumination (Fig. 1h, left) whereas activation time constants were linearly dependent on activation power on a log-log scale until the channel-intrinsic millisecond-scale activation time constant ( $\tau_{\text{on}}$ ) was approached (Fig. 1h, middle;  $n = 27$  recordings from 5 cells). The total number of photons required for photocurrents to reach a fixed fraction of  $I_{\text{max}}$  was constant over many orders of magnitude of light power (Fig. 1h, right;  $9.1 \times 10^8 \pm 1.6 \times 10^8$  photons;  $n = 27$  recordings from 5 cells). Taken together, these data indicate that SSFO acts as an integrator with photocurrent determined by total photon exposure.

## Consequences of altered E/I balance

To validate the SSFO *in vivo*, we performed optrode recordings<sup>41</sup> in anaesthetized mice expressing SSFO in prelimbic and infralimbic subregions of the mPFC. To modulate excitation, we used adeno-associated virus (AAV) expressing SSFO-EYFP (EYFP, enhanced yellow fluorescent protein) under control of the CaMKII $\alpha$  promoter (CaMKII $\alpha$ :SSFO; Supplementary Methods). To modulate inhibition we expressed a double-floxed inverted open-reading-frame (DIO) AAV<sup>32</sup> encoding SSFO-EYFP in PV::Cre (PV, parvalbumin) transgenic mice (PV::SSFO), where expression of SSFO-EYFP is restricted to GABAergic PV neurons<sup>32</sup> (Supplementary Fig. 2). Using extracellular



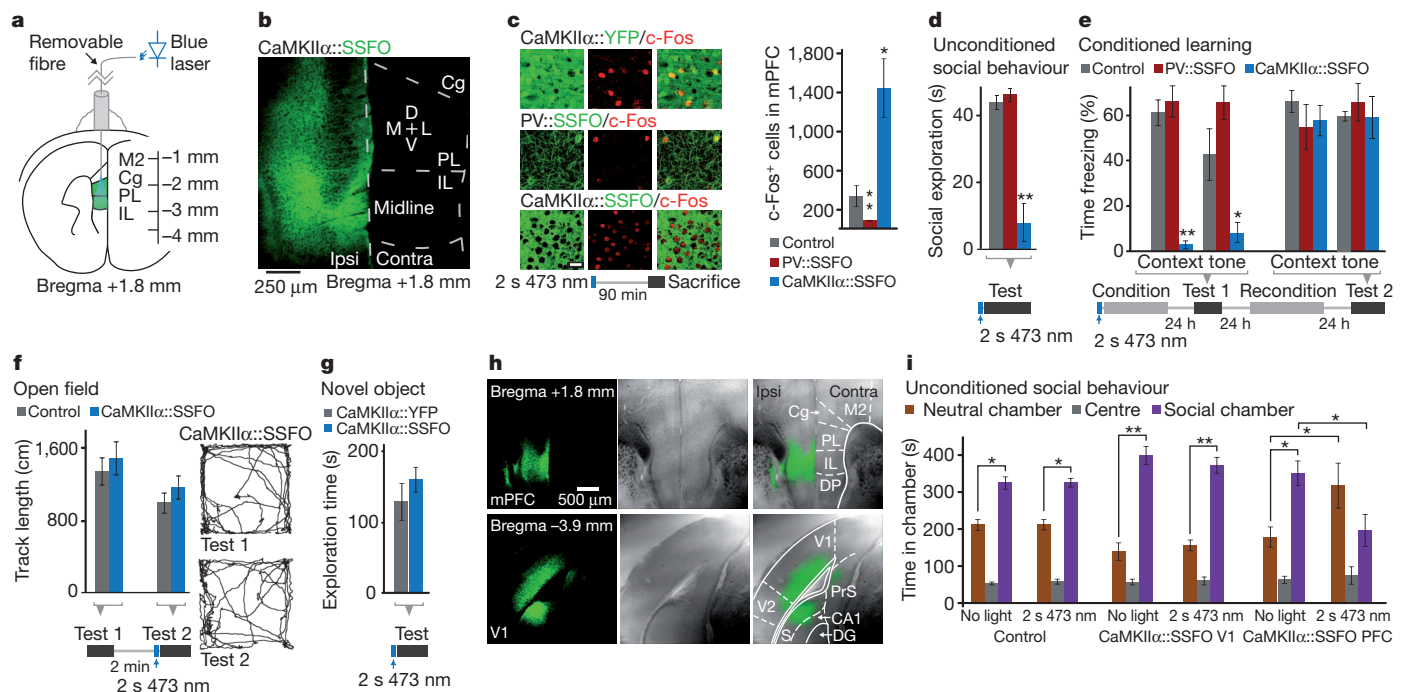
**Figure 1 | Kinetic and absorbance properties of a stabilized SFO.** **a–c**, Absorption spectra recorded after illumination with 450 nm light for 30 s.

Absorption difference spectra ( $\Delta$  Abs.) taken from the corresponding absorption spectra are shown in the insets. Spectra were collected at the indicated times after the end of illumination; note prominent recovery after 3 min in the single mutants, in contrast to the double mutant. **d**, Representative whole-cell patch-clamp recording of photocurrent in a cultured hippocampal neuron expressing ChR2(C128S/D156A) SSFO. Blue and orange bars indicate activation and deactivation light pulses. **e**, Monoexponential fits of photocurrent decay in cells expressing ChR2(C128S/D156A) (dark blue;  $\tau = 29$  min) or ChR2(D156A) (light blue;  $\tau = 6.9$  min). **f**, Activation (blue) and deactivation (orange) spectra recorded from cultured neurons expressing ChR2(C128S/D156A). **g**, Simplified photocycle scheme; in C128S/D156 mutants the transition P520 to P480 is probably slowed down or blocked, avoiding the desensitized state Des480 which cannot be reactivated with 470 nm light. Both yellow and violet light (yellow and violet arrows) shuttle the channels to an inactive state (see Supplementary Fig. 1). **h**, Whole-cell photocurrent responses of a cultured neuron expressing SSFO to 470 nm light pulses of indicated power (left). Pulse lengths were 2 s (grey traces) or 5 s (black traces). Dashed lines mark light pulse termination. Apparent time constants for activation ( $\tau_{\text{on}}$ ) are shown on a log-log plot versus light power ( $n = 27$  recordings from 5 cells; middle). Within this broad range of light power, the calculated number of incident photons arriving at each cell for photocurrents to reach the exponential curve constant (63% of  $I_{\text{max}}$ ) for that cell was constant (right). Each point represents a photon number from a single recording at a given light power.

recordings in anaesthetized mice to access this medial structure, we observed significant light-induced changes in multiunit activity restricted to transduced regions of the mPFC (Supplementary Fig. 3). In mice expressing CaMKII $\alpha$ ::SSFO, multiunit spike rates increased within the transduced area as expected (Supplementary Fig. 3b, c, e, f). Conversely, in PV::SSFO mice (despite the sparseness of PV cells and incomplete transduction; Supplementary Fig. 2), multiunit activity profoundly decreased after an identical 1 s pulse of 473 nm light and returned to near-baseline levels after a 2 s 561 nm pulse (Supplementary Fig. 3b, d, e; subsequent behavioural work described later used identical opsin-expression experimental conditions). We next examined the effects of altered E/I balance on prefrontal information transmission using whole-cell patch-clamp recordings. In the presence of ongoing synaptic activity, CaMKII $\alpha$ ::SSFO activation increased synaptic input onto both expressing and non-expressing pyramidal cells (Supplementary Fig. 4a, b). To test the input–output relationship, we stimulated CaMKII $\alpha$ ::SSFO neurons with trains of simulated excitatory post-synaptic currents (sEPSCs) designed to span a wide range of sEPSC rates over time<sup>32</sup> (Supplementary Fig. 4c and Supplementary Methods). Elevation of the cellular E/I balance in these slices by activation of CaMKII $\alpha$ ::SSFO led to saturation of the input–output (*I–O*) curve and a reduction in *I–O* mutual information (Supplementary Fig. 4d–i). In contrast, activation of SSFO in PV cells potentially reduced ongoing synaptic activity (Supplementary Fig. 5a) and decreased *I–O* curve gain in recorded pyramidal cells, as expected via synaptic inhibition.

However, the shape of the *I–O* curve was preserved with only a small reduction in dynamic range but no change in saturation and no significant effect on *I–O* mutual information (Supplementary Fig. 5b–g). Overall, the reduction in principal neuron information throughput after light activation in CaMKII $\alpha$ ::SSFO slices was 4.8-fold larger than in PV::SSFO slices ( $P = 0.0144$ ; across a wide range of analysis parameters, Supplementary Fig. 5h). These results indicate that at the principal neuron level, information processing may be more impaired by increased E/I balance.

We next examined the effects of this elevated cellular E/I balance on behaviour. We expressed SSFO in either mPFC excitatory or inhibitory PV-expressing neurons as described earlier, and implanted a chronic fibre-optic connector terminating dorsal to the mPFC (Fig. 2a, b). We confirmed, using c-Fos staining as a marker of activity, that in the awake state neural activity could be increased or decreased by activation of SSFO in pyramidal and PV neurons respectively (Fig. 2c and Supplementary Fig. 6a). In behavioural tests, all mice received a single 2 s pulse of 473 nm light through the implanted fibre-optic connector, followed by removal of the fibre optic 1 min before introduction into the behavioural chamber, capitalizing on the stability of the SSFO. Exploration of a new same-sex juvenile introduced to the experimental animal home cage was virtually abolished in the elevated E/I (CaMKII $\alpha$ ::SSFO) group following SSFO activation (Fig. 2d and Supplementary Videos 1 and 2;  $n = 8$  CaMKII $\alpha$ ::SSFO mice and  $n = 6$  controls;  $P < 0.0005$ ); PV::SSFO mice showed no effect of



**Figure 2 | Elevated, but not reduced, prefrontal E/I balance leads to behavioural impairment.** **a**, Experimental setup for behavioural experiments. Green region marks viral injection area; fibre-optic connector was attached to the light guide only transiently before testing. Cg, cingulate cortex; IL, infralimbic cortex; M2, secondary motor cortex; PL, prelimbic cortex. **b**, Confocal image from a mouse injected with CaMKII $\alpha$ ::SSFO–EYFP virus shows expression in prelimbic and infralimbic cortex. D, dorsal; L, lateral; M, medial; V, ventral. Contra, contralateral; Ipsi, ipsilateral. **c**, Representative images of prefrontal slices from PV::SSFO and CaMKII $\alpha$ ::SSFO mice stained for c-Fos 90 min after a 2 s 473 nm light pulse. Scale bar, 25  $\mu$ m. Graph shows average c-Fos-positive cell counts in the mPFC of control, CaMKII $\alpha$ ::SSFO and PV::SSFO mice. **d**, Social exploration in control, CaMKII $\alpha$ ::SSFO and PV::SSFO mice of a juvenile intruder in the home cage. CaMKII $\alpha$ ::SSFO mice showed a significant reduction in social exploration. **e**, Mice administered one 2 s 473 nm pulse of light before fear conditioning were tested the next day for freezing in response to the conditioned context or auditory cue;

CaMKII $\alpha$ ::SSFO mice were significantly impaired in freezing responses to both conditioned stimuli. On the following day, mice were reconditioned without optical stimulation and freezing was evaluated 24 h later. All mice showed similar freezing behaviour in the absence of light. **f**, Open-field exploration in CaMKII $\alpha$ ::SSFO (blue) and CaMKII $\alpha$ ::EYFP (grey) mice, before (Test 1) and after (Test 2) light activation. Example tracks from a CaMKII $\alpha$ ::SSFO mouse are shown. **g**, Novel object exploration over a 10 min period is similar in mice expressing CaMKII $\alpha$ ::SSFO (blue) and CaMKII $\alpha$ ::EYFP (grey). **h**, Phase contrast and fluorescence images of coronal sections from wild-type mice injected with CaMKII $\alpha$ ::SSFO in PFC (top) or V1 (bottom). **i**, Social behaviour in the three-chamber test is impaired following a 2 s 473 nm light pulse in mice expressing CaMKII $\alpha$ ::SSFO in the PFC ( $n = 5$ ), but not in control mice ( $n = 6$ ) or in mice expressing CaMKII $\alpha$ ::SSFO in V1 ( $n = 8$ ). Time spent in a given chamber after stimulation is only significantly altered in mice expressing CaMKII $\alpha$ ::SSFO. All bar graphs depict mean  $\pm$  s.e.m. (\* $P < 0.05$ , \*\* $P < 0.005$ ).



SSFO activation (Fig. 2d;  $n = 6$  PV::SSFO mice;  $P > 0.1$ ). To test episodic learning (also impaired in autism and schizophrenia<sup>42,43</sup>), we next subjected the same mice to a fear conditioning protocol immediately following SSFO activation. The elevated E/I (CaMKII $\alpha$ ::SSFO) mice showed no conditioned responses 24 h after conditioning (Fig. 3e; to either context,  $P < 0.0005$ ; or tone,  $P < 0.05$ , compared with controls). We then reconditioned the same mice in the absence of light activation of SSFO. Twenty-four hours later, responses of these mice were indistinguishable from controls (Fig. 2e;  $P > 0.1$  cue and context). The reduced E/I balance (PV::SSFO) group showed no significant impairment compared with controls to either tone or context (Fig. 2e;  $P = 0.09$  and  $P = 0.56$ , respectively). Elevated E/I balance-induced deficits were not a result of motor impairment as, in the same mice, open-field behaviour was normal (Fig. 2f and Supplementary Fig. 6b;  $n = 8$  CaMKII $\alpha$ ::SSFO mice and  $n = 6$  CaMKII $\alpha$ ::EYFP mice; Supplementary Fig. 8d shows unchanged immediate foot-shock response in CaMKII $\alpha$ ::SSFO mice). Additionally, CaMKII $\alpha$ ::SSFO mice showed no abnormalities in exploration of a novel object (Fig. 2g;  $n = 8$  CaMKII $\alpha$ ::SSFO and  $n = 10$  CaMKII $\alpha$ ::EYFP mice;  $P = 0.47$ ).

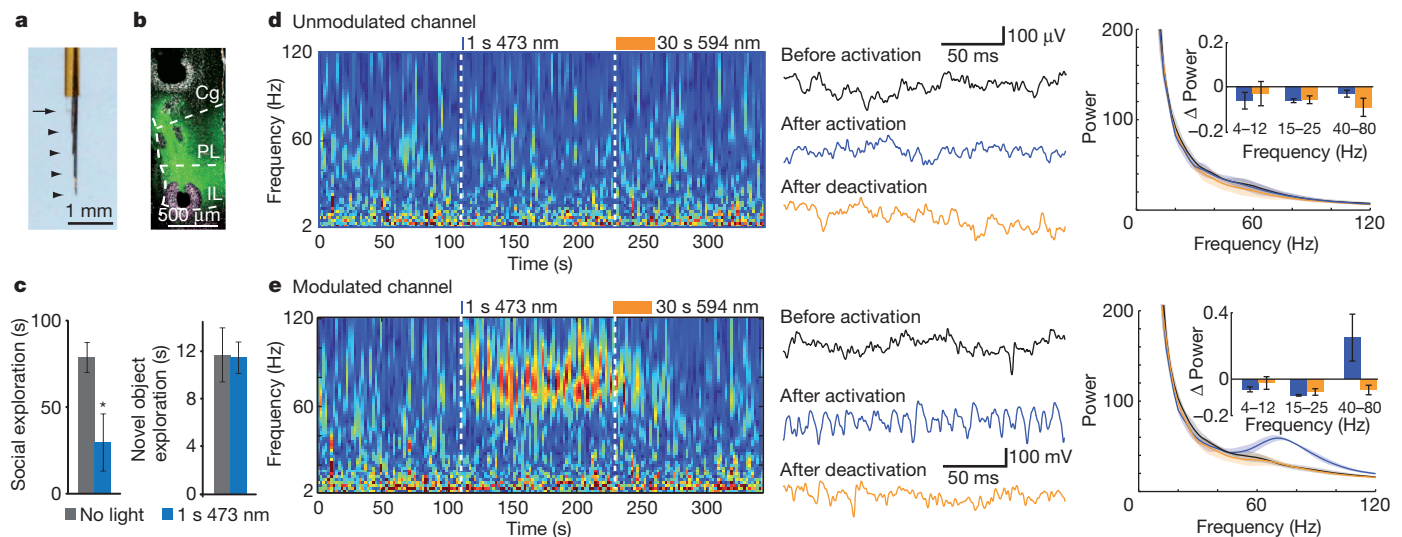
We next examined whether the social impairment induced by altered E/I balance could arise in any cortical region. We tested three groups of mice: CaMKII $\alpha$ ::SSFO virus in the mPFC, as earlier ( $n = 5$ ; Fig. 2h, top); control virus in the mPFC ( $n = 6$ ); and CaMKII $\alpha$ ::SSFO virus in the primary visual cortex (Fig. 2h, bottom; V1 area 17;  $n = 8$ ), an area in which microstimulation generates potent internal visual stimuli<sup>44</sup>. These mice were tested in the three-chamber social exploration paradigm<sup>45</sup>, an independent measure of social interaction. We found that control mice and CaMKII $\alpha$ ::SSFO V1 mice preferred the chamber with a novel conspecific mouse, in both no-light and light conditions (Fig. 2i). In contrast, CaMKII $\alpha$ ::SSFO mPFC mice exhibited normal social exploration at baseline, but significantly reduced social exploration after SSFO activation (Fig. 2i). This impairment in social behaviour was not correlated with an increased-anxiety phenotype in the elevated plus maze (Supplementary Fig. 6d). Viral targeting was evaluated post-hoc and revealed CaMKII $\alpha$ ::SSFO expression in targeted regions in all mice tested (Fig. 2h and Supplementary Fig. 7). Taken together, these experiments indicate that elevated, but not reduced, cellular E/I balance in the mPFC significantly impairs social behaviour

and conditioning, and that this deficit does not result from changes in anxiety or locomotion.

### E/I balance elevation elicits rhythmic activity

Behavioural impairment in both autism and schizophrenia has been associated with elevated baseline (non-evoked) high-frequency activity in the 30–80 Hz range<sup>20,21,46</sup>. To examine the activity patterns in awake behaving mice during E/I balance modulation, we designed a chronic multi-site optrode (CMO; Fig. 3a and Methods) to sample activity both in modulated mPFC and in unmodulated regions above and below the mPFC. We hypothesized that temporal stability and photon integration of SSFO would allow light delivery with superficial (non-brain penetrating) fibre-optic placement (Supplementary Fig. 8a, b), and verified that superficial light delivery could modulate SSFO-expressing mPFC cells using c-Fos quantification and behavioural testing (Supplementary Fig. 8c–e). We used CMOs to record neural activity in CaMKII $\alpha$ ::SSFO mice; fibre placement was superficial and electrode positions were confirmed post-hoc by electrolytic lesioning (Fig. 3b). We found that CaMKII $\alpha$ ::SSFO mice with CMO implants had normal baseline social behaviour that was impaired after a 1 s 473 nm pulse (Fig. 3c;  $n = 3$ ,  $P = 0.044$ ). Novel object exploration was unaffected by stimulation, consistent with previous findings (Fig. 3c;  $P = 0.82$ ). The activation pulse did not significantly alter locomotor behaviour in either the home cage (not shown) or a novel open field, although the latter trended towards reduced anxiety (Supplementary Fig. 8f and Supplementary Video 3).

We recorded multiunit activity and local field potentials simultaneously on all channels during these behaviour trials. In the home cage, we observed significant increases in multiunit activity after a 1 s 473 nm light pulse in channels within the SSFO-expressing regions (Supplementary Fig. 8g, h;  $77 \pm 18\%$  on modulated channels, compared with  $-3.4 \pm 4.4\%$  on unmodulated channels;  $n = 4$  modulated and 4 unmodulated channels in 3 mice averaged over 4 sweeps per mouse;  $P = 0.02$ ). Modulated channels showed a marked increase in high-frequency activity with a peak at 80 Hz after SSFO activation (Fig. 3e) delimited to the activation period (Fig. 3d, e, left). Channels with increased high-frequency activity had concomitant reduction of lower frequencies (Fig. 3e, right, inset). High-frequency activity was



**Figure 3 | Elevated cellular E/I balance in the mPFC drives baseline high frequency rhythmicity in freely moving, socially impaired mice.**

**a**, Implantable CMO for awake, behaving recordings in mouse M2 and PFC. Arrowheads indicate wire termination sites; arrow shows cleaved end of fibre-optic connector. **b**, Electrolytic lesions mark the sites from which recordings were taken in a mouse expressing CaMKII $\alpha$ ::SSFO. **c**, Social (left) and novel object exploration (right) before (grey) and after (blue) activation with 473 nm light in mice used for CMO recordings ( $n = 3$  mice). **d**, **e**, Local-field-potential

wavelet spectrogram from a representative unmodulated channel (**d**) and a representative modulated channel (**e**). Example traces are shown for the baseline, activation and deactivation periods. Average wavelet spectra for the three indicated periods ( $n = 5$  trials in 1 mouse) and population data (insets;  $n = 3$  mice) are shown, demonstrating a specific increase in gamma rhythmicity on the modulated channel after SSFO activation in PFC pyramidal neurons. All bar graphs depict mean  $\pm$  s.e.m. (\* $P < 0.05$ ).

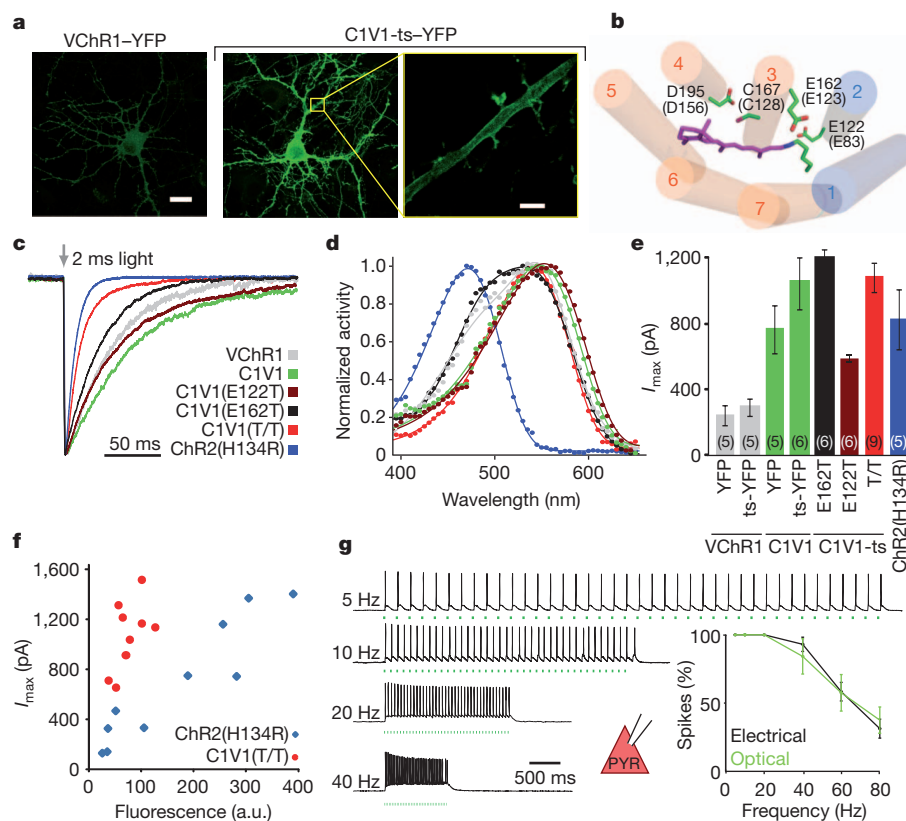
consistently increased in experiments where a 1 s 473 nm pulse was delivered during testing in the open field (Supplementary Fig. 8i and Supplementary Movie 3), during social exploration (Supplementary Fig. 8j and Supplementary Movie 5) and during novel object exploration (Supplementary Fig. 8k). Together, these data reveal that elevated rhythmic high-frequency activity, a disease correlate in clinical conditions, is elicited (along with social deficits) as a result of elevating cellular E/I balance in freely behaving mammals.

### Development of potent redshifted control

We next sought to complement SSFO with a redshifted, fast-kinetic-response excitatory ChR for combinatorial control in the mPFC. Neurons expressing VChR1 showed weak expression (Fig. 4a) and small photocurrents (Supplementary Fig. 9a), consistent with previous findings<sup>39</sup>. Adding a Kir2.1 trafficking signal<sup>47</sup> (ts) to generate VChR1-ts-EYFP only modestly enhanced photocurrents (Supplementary Fig. 9a). We next systematically produced ChR1/VChR1 chimaeras and screened for enhanced membrane expression and photocurrent in HEK cells. We observed markedly improved membrane expression when VChR1 helices 1 and 2 were replaced with ChR1 homologues (Supplementary Fig. 9b) and selected two of these (fused at ChR1 Ala 145 or Trp 163) for further characterization (Supplementary Fig. 9a, c, d). In cultured neurons, the latter variant expressed more robustly and trafficked well to the membrane (Fig. 4a and Supplementary Fig. 9e, f) with enhanced peak photocurrents ( $888 \pm 128$  pA,  $n = 11$  cells;

$P < 0.0005$ ) compared with VChR1 (Supplementary Fig. 9a). The action spectrum remained redshifted (Supplementary Table 1 and Fig. 4d), while ion selectivity was similar to ChR2 and VChR1 (Supplementary Fig. 9g). The Kir2.1 trafficking sequence<sup>47</sup> further increased photocurrents ( $1104 \pm 123$  pA,  $n = 12$  cells; Fig. 4e, Supplementary Fig. 9a and Supplementary Tables 1, 2). This ChR1/VChR1 chimaera contains no ChR2 sequence at all and is referred to here as C1V1.

Fast deactivation<sup>37</sup> of a redshifted opsin would be optimal for temporal and spectral separation from SSFO. However, C1V1-ts-EYFP photocurrents decayed >10-fold more slowly than those of ChR2, and indeed more slowly than those of VChR1 (Fig. 4c;  $\tau_{\text{off}}$   $156 \pm 12$  ms and  $132 \pm 12$  ms for C1V1-ts-EYFP ( $n = 4$ ) and VChR1-EYFP ( $n = 5$ ), respectively; Supplementary Table 1). To enhance photocurrent kinetics of C1V1, we used structural models to search the chromophore region<sup>33,37</sup> (Fig. 4b) for residues that might modulate photocycle kinetics, inactivation and blue absorption. We found that the E162T mutation (homologous to a ChETA mutation in ChR2) and the E122T mutation both significantly reduced  $\tau_{\text{off}}$  (Fig. 4c). The combined E122T/E162T mutation synergistically enhanced photocurrent temporal precision while maintaining robust photocurrent sizes and reducing the residual shorter-wavelength response observed in VChR1 and C1V1(E162T) (Fig. 4c–e). Photocurrent sizes generally correlated with fluorescence intensity across constructs (Supplementary Fig. 9h), but C1V1(E122T/E162T) cells showed stronger photocurrents than ChR2(H134R) cells at equivalent fluorescence levels (Fig. 4f).



**Figure 4 | Multistep engineering of a potent redshifted ChR.** **a**, Confocal images of cultured hippocampal neurons transfected with VChR1-EYFP or C1V1-ts-EYFP under the control of the CaMKII $\alpha$  promoter. Yellow box denotes region expanded in the last panel, showing dendritic membrane localization of C1V1-ts-EYFP. Scale bars: 20  $\mu$ m (left, middle), 4  $\mu$ m (right). **b**, Model of the C1V1 chromophore binding pocket, showing ChR1 helices in blue, VChR1 helices in orange, the retinal Schiff base (RSB) in purple, and key amino acid residues (with corresponding ChR2 numbering in parentheses and the modelled location of the SSFO mutations C128 and D156 shown for context). **c**, Representative traces and summary plot of channel closure time constant ( $\tau_{\text{off}}$ ) in cultured neurons expressing the indicated channelrhodopsins; traces are normalized to peak current. **d**, Action spectra collected for the

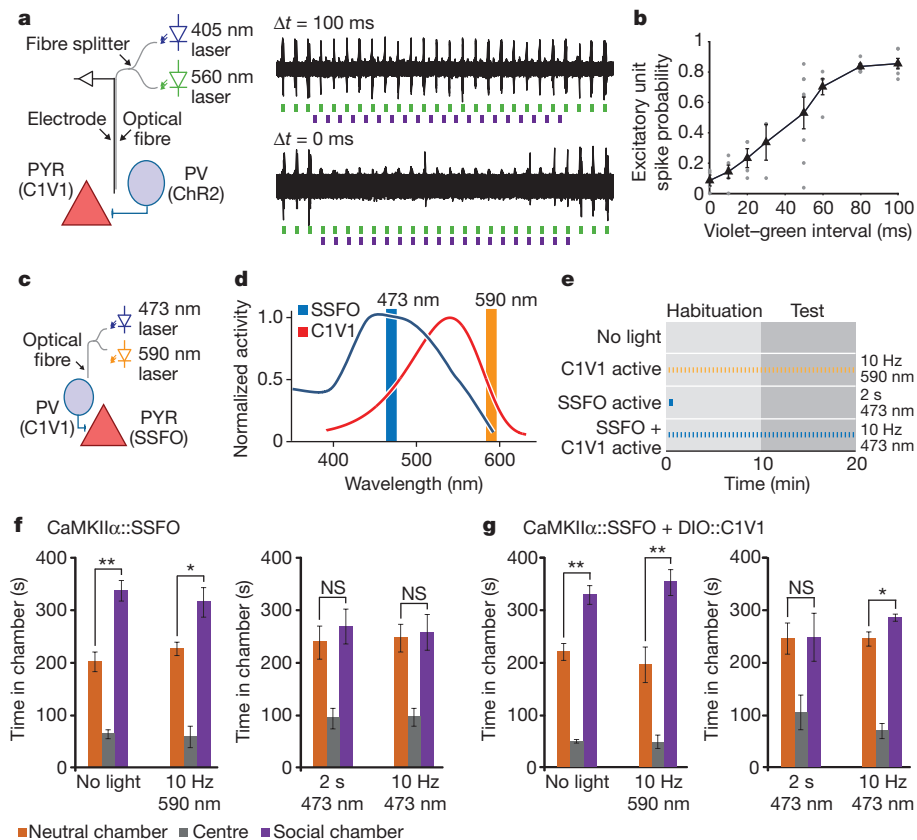
indicated channelrhodopsins (colour code as in c). Photocurrents were collected with 2 ms light pulses in HEK293 cells. **e**, Mean peak photocurrents recorded in cultured neurons expressing the indicated channelrhodopsins in response to a 2 ms 542 nm light pulse. Colours are as indicated in c; numbers in brackets indicate  $n$ . **f**, Fluorescence-photocurrent relationship in ChR2(H134R) (blue) and C1V1(E122T/E162T) (red). a.u., arbitrary units. **g**, Acute slice recordings in prefrontal pyramidal neurons (PYR) expressing C1V1(E122T/E162T) and stimulated with 560 nm light pulse trains or current injections at the indicated frequencies. Summary graphs show light and current-evoked spike probability versus stimulation frequency ( $n = 6$ ;  $P > 0.4$  at all frequencies). All error bars indicate s.e.m.

Next we injected mice with AAV5-CaMKII $\alpha$ ::C1V1(E122T/E162T)-ts-EYFP virus in the mPFC and recorded from expressing neurons in acute slices. The frequency responses to trains of 2 ms light pulses or current injection pulses were indistinguishable (Fig. 4g;  $n = 6$  cells in 2 slices), suggesting that in this case properties and not opsin kinetics limit performance (Supplementary Fig. 10). We next found that with C1V1(E122T) (the most redshifted of all variants although somewhat slower than C1V1(E122T/E162T); Fig. 4d) reliable spiking with 630 nm light could be obtained (Supplementary Fig. 11). Lastly, using AAV5-CaMKII $\alpha$ ::C1V1(E122T/E162T) together with AAV5-EF1 $\alpha$ ::ChR2(H134R), we were able to differentially drive spiking in cortical pyramidal and PV neurons with 560 nm and 405 nm light, respectively (Supplementary Fig. 12). Combinatorial excitation was also possible with SSFO and C1V1, as peak frequencies of evoked rhythmic circuit activity generated in prefrontal cortical slices by C1V1 activation seemed to be shifted to higher frequencies by co-activation of SSFO in PV neurons (Supplementary Fig. 13). Furthermore, using C1V1 and ChR2 we were able to independently activate axons of two converging excitatory pathways (cortico-thalamic and thalamo-cortical) onto thalamic reticular nucleus neurons, observing that under these conditions the temporal coincidence of subthreshold synaptic inputs from each pathway can evoke spiking in target

neurons (Supplementary Fig. 14). Together these results indicated that combinatorial control within intact tissue is enabled with these C1V1 variants.

### Increased inhibition modulates behavioural deficits

In principle, an elevated E/I ratio could be achieved not only by increasing excitation but also by reducing inhibition, although this loss-of-function approach might be operative only under high baseline activity of the inhibitory cells (Supplementary Figs 6c, 15). However, a question salient to the elevated cellular E/I ratio hypothesis is whether adding an additional disruption in the form of increased inhibition could ameliorate behavioural deficits associated with elevated E/I balance. For this experiment, we first had to verify that combinatorial excitation could be feasible *in vivo*. We therefore explored combinatorial excitation of mPFC pyramidal and PV neurons *in vivo* with optrode recordings (Fig. 5a, left). To examine inhibitory effects of PV cell activity on pyramidal neuron spiking *in vivo*, we expressed CaMKII $\alpha$ ::C1V1(E122T/E162T) and DIO::ChR2(H134R) in the mPFC of PV::Cre mice. Using an optrode, we applied 5 Hz 405 nm pulses (to activate ChR2 in PV cells) followed by 5 Hz 561 nm pulses (to activate C1V1 in pyramidal neurons) while varying inter-pulse interval (Fig. 5a, right). The fidelity of green-light-evoked



**Figure 5 | Combinatorial optogenetics enables partial reversal of elevated E/I-balance social behaviour disruption.** **a**, mPFC optrode recording in an anaesthetized PV::Cre mouse injected with CaMKII $\alpha$ ::C1V1(E162T)-ts-EYFP and Efla-DIO::ChR2-EYFP (diagram illustrates experimental setup). Violet (405 nm) light pulses are presented with variable delay ( $\Delta t$ ) relative to green light pulses (example traces). **b**, Summary graph shows probability of green-light-evoked spikes with violet pulses preceding the green light pulses by the indicated delays. Individual points are from single recordings. Black line shows average for all recordings ( $>3$  recording sites per bin). **c**, Experimental paradigm for SSFO activation in pyramidal neurons and C1V1 activation in PV neurons. **d**, Action spectra of SSFO (blue) and C1V1(E122T/E162T) (C1V1, red). Orange and blue vertical lines indicate stimulation wavelengths used in the experiments. **e**, Experiment design and pulse patterns; no-light control was

used for baseline behaviour; 2 s 473 nm light for prolonged SSFO activation; 10 Hz 473 nm for co-activation of SSFO and C1V1; 10 Hz 590 nm for C1V1 activation. **f**, Mice expressing CaMKII $\alpha$ ::SSFO ( $n = 7$ ) showed significant social preference at baseline, but exhibited social dysfunction after either 2 s 473 nm activation or during 10 Hz 473 nm activation. **g**, Mice expressing both CaMKII $\alpha$ ::SSFO and DIO-PV::C1V1 ( $n = 7$ ) showed impaired social behaviour after a 2 s 473 nm pulse, but showed partially restored social behaviour during the 10 Hz 473 nm light stimulation. Activation of C1V1 alone with 10 Hz 590 nm pulses did not impair social behaviour. NS, not significant. Supplementary Fig. 16 shows normal social behaviour under all illumination conditions in YFP-expressing control cohorts. All error bars indicate s.e.m. ( $*P < 0.05$ ;  $**P < 0.005$ ).



events dropped as the inter-pulse interval decreased, with maximal excitatory responses observed with an interval of 100 ms (Fig. 5a, top trace) and responses were nearly abolished when light pulses were presented with sufficient synchrony (Fig. 5a, bottom trace; summary data in Fig. 5b). These data demonstrate combinatorial optogenetic control within an intact mammal.

We next leveraged combinatorial optogenetics *in vivo* to excite mPFC pyramidal cells with SSFO and PV neurons with C1V1 (E122T/E162T) in freely moving mice (Fig. 5c; referred to here as SSFO/C1V1 mice;  $n = 7$ ). Control cohorts were injected with CaMKII $\alpha$ ::SSFO virus alone ( $n = 7$ ) or with CaMKII $\alpha$ ::EYFP ( $n = 8$ ). Mice were tested in the three-chamber social test under four different illumination paradigms, designed to examine behaviour with baseline properties, elevated excitatory tone, elevated inhibitory tone and co-elevated excitatory and inhibitory tone (Fig. 5d, e). All mice showed significant preference for the social chamber during baseline testing (Fig. 5f, g and Supplementary Fig. 16; CaMKII $\alpha$ ::SSFO mice,  $P = 0.0002$ ; SSFO/C1V1 mice,  $P = 0.0014$ ; CaMKII $\alpha$ ::EYFP mice,  $P = 0.0029$ ). Under pulsed laser light at 590 nm, all mice showed significant preference for the social chamber (Fig. 5f, g and Supplementary Fig. 16; CaMKII $\alpha$ ::SSFO mice,  $P = 0.02$ ; SSFO/C1V1 mice,  $P = 0.0048$ ; CaMKII $\alpha$ ::EYFP mice,  $P = 0.039$ ). After a 2 s 473 nm pulse given at the beginning of the test, neither the CaMKII $\alpha$ ::SSFO nor the SSFO/C1V1 mice preferred the social chamber (Fig. 5f, g and Supplementary Fig. 16;  $P = 0.52$  and  $P = 0.96$ , respectively). Lastly, we sought to ameliorate the behaviour deficit by compensating cellular E/I balance, co-activating SSFO in excitatory cells and C1V1 (E122T/E162T) in inhibitory cells by pulsing 10 Hz 473 nm light throughout the test period (Fig. 5f, g). CaMKII $\alpha$ ::SSFO mice (with no C1V1 (E122T/E162T)) and thus experiencing purely cellular E/I balance elevation) showed no significant preference for the social chamber (Fig. 5f;  $P = 0.81$ ); in contrast, despite the sparseness of the PV population and incomplete transduction under these conditions (Supplementary Fig. 2) SSFO/C1V1 showed modest but significant preference for the social chamber (Fig. 5g;  $P = 0.026$ ). Control CaMKII $\alpha$ ::EYFP mice significantly preferred the social chamber under both the 2 s 473 nm and the 10 Hz 473 nm paradigms (Supplementary Fig. 16). Together, this series of experiments indicate that social deficits seen with elevated mPFC cellular E/I balance may be partially reduced by restoring balance with increased inhibitory tone.

## Discussion

Several lines of evidence have indicated the involvement of elevated cellular E/I balance in the aetiology of medication-unresponsive social and information-processing impairments in autism and schizophrenia, although it has been unclear if such an effect would operate on the acute or chronic timescale<sup>3,9,10,23,28,48</sup>. Moreover, the tight interplay and complexity of excitation and inhibition within cortical microcircuitry have precluded the direct investigation of cell- and circuit-level effects of changes in cellular E/I balance; for example, cell-specific pharmacological agents are lacking, and homeostatic processes can occur downstream of synaptic and intrinsic excitability alterations. We therefore sought to investigate the effects of this altered cellular E/I balance on the acute timescale and quantify the effects on information transmission, network activity and animal behaviour. To stably modulate neural activity in awake animals, we produced an SSFO that enabled combined spectral/temporal separation when used together with red-shifted opsins (C1V1 variants), developed as chimaeras along with combinatorial mutagenesis of the chromophore region to enhance potency, kinetics and spectral properties.

With these new tools, we obtained causal support for the elevated cellular E/I balance hypothesis, and identified circuit-physiology manifestations of the resulting social dysfunction. SSFO activation in prefrontal cortical excitatory (but not inhibitory) neurons led to profound, yet reversible, impairments in social function and cognition without motor abnormalities or increased anxiety. To understand more fully the elevated E/I state, we probed both *in vitro* and *in vivo* the

underlying circuit-physiology manifestations. Despite the complexity resulting from the broad range of circuit phenomena that a cellular E/I balance modulation could initiate, we found that cellular E/I balance elevation, but not reduction under these conditions, altered the transfer functions of principal neurons in a way that quantitatively impaired information transmission within cortical circuitry. The effects of elevated E/I balance on social behaviour showed evidence of specificity for the PFC, as increasing the E/I ratio in primary visual cortex did not impair social behaviour. The PFC network, with its extensive sub-cortical connectivity, might therefore be particularly susceptible to eliciting psychiatric symptom-related effects in the setting of subtle changes in E/I balance, a notion that is supported by observed alterations in PFC inhibitory markers associated with psychiatric disease<sup>12,13,48–50</sup> and the altered PFC rhythmicity observed in autistic individuals<sup>20</sup>. And using spectrottemporal separation of the activity of SSFO and C1V1, we found that increased cellular inhibition moderately ameliorated the social behaviour deficits in mice that had been subjected to elevation of cellular E/I balance.

We also identified correspondence between a clinical marker of disease states linked to social dysfunction (elevated baseline gamma power) and electrophysiological findings during free behaviour in the elevated cellular E/I state. We found that the elevated E/I state was associated with robust, stable high-frequency power in the 30–80 Hz range generated by and manifested within the regions directly experiencing elevated cellular E/I balance. Stimulus-evoked gamma oscillations had previously been shown to be important for refining information flow through cortical circuitry<sup>32</sup>, but high baseline (unevoked) gamma oscillations may interfere with cortical function and contribute to disease states<sup>20,21</sup>. Together, these temporally delimited, cell-type-specific optogenetic manipulations (although fully capable of causing either acute or chronic changes in complex local circuit processing and input–output relationships) distinguish direct real-time effects of cellular E/I balance changes *in vivo* from the many possible developmental and structural alterations that could occur with more chronic interventions<sup>25–29</sup>. Despite the complexity of homeostatic processes that contribute to behavioural phenotypes in genetic models, our data are consistent with the E/I balance hypothesis and with recent reports of impaired social function and learning in mouse models of Rett syndrome<sup>28,29</sup>. In sum, these findings may represent a step towards understanding the pathophysiology of social and information-processing dysfunction, and despite the complexity of cortical processing, together provide causal support for the cellular E/I balance elevation hypothesis of severe psychiatric disease symptomatology on the cellular, circuit and behavioural levels.

## METHODS SUMMARY

C1V1 was cloned by overlap extension PCR, fusing the N-terminal sequence of the *Chlamydomonas reinhardtii* ChR1 coding sequence<sup>40</sup> with the C-terminal sequence of *Volvox carterii* ChR1<sup>39</sup>. SSFO mutants were generated by site-directed mutagenesis of the humanized ChR2 sequence. Constructs were cloned into the pLenti or pAAV-MCS backbone under control of the CaMKII $\alpha$  promoter or using a double-floxed inverted open-reading frame (ORF) Cre-dependent backbone<sup>32</sup>. Maps and clones are available at <http://www.optogenetics.org>.

Virus was stereotactically injected 2–3 weeks before physiological experiments or behavioural testing. Optrode recordings in anaesthetized mice were conducted as described<sup>41</sup> using a dual-laser system with either 473 nm and 561 nm lasers for SSFO experiments or 405 nm and 561 nm lasers for combinatorial excitation experiments with ChR2 and C1V1.

Recordings in behaving mice were performed using custom-built 4-microwire arrays attached to an optical fibre. Arrays were designed to sample from a large anatomical volume including transduced and untransduced tissue. Data were analysed in Matlab, using custom software; spectral analysis was performed using the wavelet method<sup>32</sup>.

Received 17 September 2010; accepted 10 July 2011.

Published online 27 July 2011.

1. Pardo, C. A. & Eberhart, C. G. The neurobiology of autism. *Brain Pathol.* **17**, 434–447 (2007).

2. O'Donovan, M. C., Craddock, N. J. & Owen, M. J. Genetics of psychosis; insights from views across the genome. *Hum. Genet.* **126**, 3–12 (2009).
3. Südhof, T. C. Neuroligins and neuroligins link synaptic function to cognitive disease. *Nature* **455**, 903–911 (2008).
4. Patterson, P. H. Modeling autistic features in animals. *Pediatr. Res.* **69**, 34R–40R (2011).
5. Walsh, T. *et al.* Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* **320**, 539–543 (2008).
6. Markram, K. & Markram, H. The intense world theory—a unifying theory of the neurobiology of autism. *Front. Hum. Neurosci.* **4**, 224 (2010).
7. Vattikuti, S. & Chow, C. C. A computational model for cerebral cortical dysfunction in autism spectrum disorders. *Biol. Psychiatry* **67**, 672–678 (2010).
8. Kehr, C., Maziashvili, N., Dugladze, T. & Gloveli, T. Altered excitatory-inhibitory balance in the NMDA-hypofunction model of schizophrenia. *Front. Mol. Neurosci.* **1**, 6 (2008).
9. Rubenstein, J. L. Three hypotheses for developmental defects that may underlie some forms of autism spectrum disorder. *Curr. Opin. Neurol.* **23**, 118–123 (2010).
10. Rubenstein, J. L. & Merzenich, M. M. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain Behav.* **2**, 255–267 (2003).
11. Gogolla, N. *et al.* Common circuit defect of excitatory-inhibitory balance in mouse models of autism. *J. Neurodevel. Disord.* **1**, 172–181 (2009).
12. Hashimoto, T. *et al.* Conserved regional patterns of GABA-related transcript expression in the neocortex of subjects with schizophrenia. *Am. J. Psychiatry* **165**, 479–489 (2008).
13. Hashimoto, T. *et al.* Gene expression deficits in a subclass of GABA neurons in the prefrontal cortex of subjects with schizophrenia. *J. Neurosci.* **23**, 6315–6326 (2003).
14. Belforte, J. E. *et al.* Postnatal NMDA receptor ablation in corticolimbic interneurons confers schizophrenia-like phenotypes. *Nature Neurosci.* **13**, 76–83 (2010).
15. Blatt, G. J. *et al.* Density and distribution of hippocampal neurotransmitter receptors in autism: an autoradiographic study. *J. Autism Dev. Disord.* **31**, 537–543 (2001).
16. Bourgeron, T. A synaptic trek to autism. *Curr. Opin. Neurobiol.* **19**, 231–234 (2009).
17. Belmonte, M. K., Gomot, M. & Baron-Cohen, S. Visual attention in autism families: 'unaffected' sibs share atypical frontal activation. *J. Child Psychol. Psychiatry* **51**, 259–276 (2010).
18. Gomot, M., Belmonte, M. K., Bullmore, E. T., Bernard, F. A. & Baron-Cohen, S. Brain hyper-reactivity to auditory novel targets in children with high-functioning autism. *Brain* **131**, 2479–2488 (2008).
19. Dichter, G. S., Felder, J. N. & Bodfish, J. W. Autism is characterized by dorsal anterior cingulate hyperactivation during social target detection. *Soc. Cogn. Affect. Neurosci.* **4**, 215–226 (2009).
20. Orekhova, E. V. *et al.* Excess of high frequency electroencephalogram oscillations in boys with autism. *Biol. Psychiatry* **62**, 1022–1029 (2007).
21. Rojas, D. C., Maharajh, K., Teale, P. & Rogers, S. J. Reduced neural synchronization of gamma-band MEG oscillations in first-degree relatives of children with autism. *BMC Psychiatry* **8**, 66 (2008).
22. Gillberg, C. & Billstedt, E. Autism and Asperger syndrome: coexistence with other clinical disorders. *Acta Psychiatr. Scand.* **102**, 321–330 (2000).
23. Canitano, R. Epilepsy in autism spectrum disorders. *Eur. Child Adolesc. Psychiatry* **16**, 61–66 (2007).
24. Rippon, G., Brock, J., Brown, C. & Boucher, J. Disordered connectivity in the autistic brain: challenges for the “new psychophysiology”. *Int. J. Psychophysiol.* **63**, 164–172 (2007).
25. Dani, V. S. *et al.* Reduced cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett syndrome. *Proc. Natl Acad. Sci. USA* **102**, 12560–12565 (2005).
26. Etherton, M. R., Blaiss, C. A., Powell, C. M. & Südhof, T. C. Mouse neuroligin-1 $\alpha$  deletion causes correlated electrophysiological and behavioral changes consistent with cognitive impairments. *Proc. Natl Acad. Sci. USA* **106**, 17998–18003 (2009).
27. Tabuchi, K. *et al.* A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science* **318**, 71–76 (2007).
28. Chao, H. T. *et al.* Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature* **468**, 263–269 (2010).
29. Moretti, P. *et al.* Learning and memory and synaptic plasticity are impaired in a mouse model of Rett syndrome. *J. Neurosci.* **26**, 319–327 (2006).
30. Rinaldi, T., Perrodin, C. & Markram, H. Hyper-connectivity and hyper-plasticity in the medial prefrontal cortex in the valproic acid animal model of autism. *Front. Neural Circuits* **2**, 4 (2008).
31. Adamantidis, A. R., Zhang, F., Aravanis, A. M., Deisseroth, K. & de Lecea, L. Neural substrates of awakening probed with optogenetic control of hypocretin neurons. *Nature* **450**, 420–424 (2007).
32. Sohal, V. S., Zhang, F., Yizhar, O. & Deisseroth, K. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* **459**, 698–702 (2009).
33. Berndt, A., Yizhar, O., Gunaydin, L. A., Hegemann, P. & Deisseroth, K. Bi-stable neural state switches. *Nature Neurosci.* **12**, 229–234 (2009).
34. Diester, I. *et al.* An optogenetic toolbox designed for primates. *Nature Neurosci.* **14**, 387–397 (2011).
35. Bamann, C., Gueta, R., Kleinlogel, S., Nagel, G. & Bamberg, E. Structural guidance of the photocycle of channelrhodopsin-2 by an interhelical hydrogen bond. *Biochemistry* **49**, 267–278 (2010).
36. Yizhar, O., Fenno, L., Davidson, T. J., Mogri, M. & Deisseroth, K. Optogenetics in neural systems. *Neuron* **71**, 9–34 (2011).
37. Gunaydin, L. A. *et al.* Ultrafast optogenetic control. *Nature Neurosci.* **13**, 387–392 (2010).
38. Wen, L. *et al.* Opto-current-clamp actuation of cortical neurons using a strategically designed channelrhodopsin. *PLoS ONE* **5**, e12893 (2010).
39. Zhang, F. *et al.* Red-shifted optogenetic excitation: a tool for fast neural control derived from *Volvox carter*. *Nature Neurosci.* **11**, 631–633 (2008).
40. Nagel, G. *et al.* Channelrhodopsin-1: a light-gated proton channel in green algae. *Science* **296**, 2395–2398 (2002).
41. Gradinaru, V. *et al.* Targeting and readout strategies for fast optical neural control *in vitro* and *in vivo*. *J. Neurosci.* **27**, 14231–14238 (2007).
42. Lind, S. E. & Bowler, D. M. Episodic memory and episodic future thinking in adults with autism. *J. Abnorm. Psychol.* **119**, 896–905 (2010).
43. D'Argembeau, A., Raffard, S. & Van der Linden, M. Remembering the past and imagining the future in schizophrenia. *J. Abnorm. Psychol.* **117**, 247–251 (2008).
44. Ni, A. M. & Maunsell, J. H. Microstimulation reveals limits in detecting different signals from a local cortical region. *Curr. Biol.* **20**, 824–828 (2010).
45. Moy, S. S. *et al.* Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. *Genes Brain Behav.* **3**, 287–302 (2004).
46. Wilson, T. W., Rojas, D. C., Reite, M. L., Teale, P. D. & Rogers, S. J. Children and adolescents with autism exhibit reduced MEG steady-state gamma responses. *Biol. Psychiatry* **62**, 192–197 (2007).
47. Gradinaru, V. *et al.* Molecular and cellular approaches for diversifying and extending optogenetics. *Cell* **141**, 154–165 (2010).
48. Lewis, D. A., Hashimoto, T. & Volk, D. W. Cortical inhibitory neurons and schizophrenia. *Nature Rev. Neurosci.* **6**, 312–324 (2005).
49. Lewis, D. A., Volk, D. W. & Hashimoto, T. Selective alterations in prefrontal cortical GABA neurotransmission in schizophrenia: a novel target for the treatment of working memory dysfunction. *Psychopharmacology (Berl.)* **174**, 143–150 (2004).
50. Lisman, J. E. *et al.* Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. *Trends Neurosci.* **31**, 234–242 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the K.D., P.H. and J.R.H. laboratories for discussions on the manuscript. We are grateful to S. Pak, Z. Chen and C. Perry for technical assistance. O.Y. is supported by the Human Frontier Science Program. L.E.F. is supported by the Stanford MSTP program. P.H. is supported by the DFG (HE3824/9-1 and 17-1, Cluster of Excellence: Unifying Concepts in Catalysis), and K.D. by NIMH, NIDA, NINDS, the DARPA REPAIR program, CIRM and the Yu, Woo, Snyder and Keck Foundations.

**Author Contributions** O.Y., M.P., F.S. and C.R. designed and cloned all DNA constructs; O.Y. and L.E.F. contributed to all neuronal electrophysiology and behavioural experiments; T.J.D. designed the CMO implant; I.G. and J.F. contributed to behaviour and histology experiments; D.J.O. and V.S.S. contributed to slice electrophysiology and mutual information analysis; K.S. and R.F. performed spectroscopy experiments; M.P. and F.S. performed HEK cell experiments; P.H. analysed and supervised spectroscopy and HEK cell work; J.T.P. and J.R.H. conducted and analysed, and J.R.H. supervised, thalamic slice experiments. K.D. supervised all aspects of the project and O.Y., L.E.F. and K.D. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to O.Y. (ofer.yizhar@weizmann.ac.il), P.H. (hegemape@rz.hu-berlin.de) or K.D. (deissero@stanford.edu).